Exercises to “The Theory and Applications of Benford’s Law”

Steven J. Miller (editor)

PRINCETON UNIVERSITY PRESS
PRINCETON AND OXFORD

Contents

Preface v

Notation vii

PART 1. GENERAL THEORY I: BASIS OF BENFORD'S LAW 1

Chapter 1. A Quick Introduction to Benford's Law 3

Chapter 3. Fourier Analysis and Benford's Law 7

3.1 Problems from Introduction to Fourier Analysis 7
3.2 Problems from Chapter 1: Revisited 13
3.3 Problems from Chapter 3 13

PART 2. GENERAL THEORY II: DISTRIBUTIONS AND RATES OF CONVERGENCE 17

Chapter 4. Benford's Law Geometry 19

Chapter 5. Explicit Error Bounds via Total Variation 20

Chapter 6. Lévy Processes and BenfordŠs Law 21

PART 3. APPLICATIONS I: ACCOUNTING AND VOTE FRAUD 27

Chapter 7. Benford's Law as a Bridge between Statistics and Accounting 29

Chapter 8. Detecting Fraud and Errors Using Benford's Law 31

Chapter 9. Can Vote Counts' Digits and Benford's Law Diagnose Elections? 32

Chapter 10. Complementing Benford's Law for small N: a local bootstrap 33

PART 4. APPLICATIONS II: ECONOMICS 35

Chapter 11. Measuring the Quality of European Statistics 37

Chapter 12. Benford's Law and Fraud in Economic Research 38

Chapter 13. Testing for Strategic Manipulation of Economic and Financial Data 40

PART 5. APPLICATIONS III: PSYCHOLOGY AND THE SCIENCES 41

Chapter 14. Psychology and Benford's Law 43

Chapter 15. Managing Risk in Numbers Games: Benford's Law and the Small Number Phenomenon 44

Chapter 16. Benford's Law in the Natural Sciences 45

Chapter 17. Generalizing Benford's Law: A Re-examination of Falsified Clinical Data 46

PART 6. APPLICATIONS IV: IMAGES 47

Chapter 18. Partial Volume Modeling of Medical Imaging Systems using the Benford Distribution 49

Chapter 19. Application of Benford's Law to Images 50

Preface

As Benford's law arises in so many different fields, there is a tremendous opportunity to design an interdisciplinary class on its theory and applications. Our hope is that this collection of exercises for the book *The Theory and Applications of Benford's Law* will facilitate its use as a textbook.

There are almost as many possible classes that could be offered on Benford's law as there are researches in the subject; the goal of our book and these exercises is to provide a good, common denominator starting ground. As the book begins with several chapters where the theory is developed to various levels, from an informal discussion to a detailed discourse involved advanced analysis, one option is to use these chapters as a springboard to motivate numerous topics in advanced analysis.

After the theory, we turn to some of the large number of applications. The instructor has many options here, and can use these chapters and exercises to illustrate the far-reaching consequences of simple ideas. Additionally, this material may be used to motivate a quick introduction to statistics.

In order to assist instructors and students, for many chapters at

http://web.williams.edu/Mathematics/sjmiller/public_html/benford/

we provide links to relevant material and other resources (such as videos and programs) as available. Many authors contributed exercises for their and other chapters. Additionally, numerous other problems were written or assembled by the editor and students working with him (especially John Bihn of Williams College). It is a pleasure to thank everyone for their contributions. Further, one advantage of posting problems online is that this need not be a static list, and thus please feel free to email the editor suggestions for additional exercises.

Finally, if you are interested in using this book for a class, or have done so and have suggestions or requests, we would love to hear from you and work with you. Please contact the editor at the email address below.

Steven J. Miller
Williams College
Williamstown, MA
June 2014
sjm1@williams.edu, Steven.Miller.MC.96@aya.yale.edu

Notation

$\square$: indicates the end of a proof.

$\equiv$: $x \equiv y \mod n$ means there exists an integer a such that $x = y + an$.

$\exists$: there exists.

$\forall$: for all.

$|\cdot|$: $|S|$ (or $\#S$) is the number of elements in the set S.

$\lceil\cdot\rceil$: $\lceil x\rceil$ is the smallest integer greater than or equal to x, read "the ceiling of x."

$\lfloor\cdot\rfloor$ or $[\cdot]$: $\lfloor x\rfloor$ (also written $[x]$) is the greatest integer less than or equal to x, read "the floor of x."

$\{\cdot\}$ or $\langle\cdot\rangle$: $\{x\}$ is the fractional part of x; note $x = [x] + \{x\}$.

$\ll, \gg$: see Big-Oh notation.

$\vee$: $a \vee b$ is the maximum of a and b.

$\wedge$: $a \wedge b$ is the minimum of a and b.

$\mathbb{1}_A$ (or I_A) : the indicator function of set A; thus $\mathbb{1}_A(x)$ is 1 if $x \in A$ and 0 otherwise.

δ_a : Dirac probability measure concentrated at $a \in \Omega$.

λ : Lebesgue measure on $(\mathbb{R}, \mathcal{B})$ or parts thereof.

$\lambda_{a,b}$: normalized Lebesgue measure (uniform distribution) on $\big([a, b), \mathcal{B}[a, b)\big)$.

$\sigma(f)$: the σ-algebra generated by the function $f : \Omega \to \mathbb{R}$.

$\sigma(A)$: the spectrum (set of eigenvalues) of a $d \times d$-matrix A.

A^c : the complement of A in some ambient space Ω clear from the context; i.e., $A^c = \{\omega \in \Omega : \omega \notin A\}$.

$A\backslash B$: the set of elements of A not in B, i.e., $A\backslash B = A \cap B^c$.

$A\Delta B$: the symmetric difference of A and B, i.e., $A\Delta B = A\backslash B \cup B\backslash A$.

a.e. : (Lebesgue) almost every (or almost everywhere).

a.s. : almost surely, i.e., with probability one.

$\mathbb{B}$: Benford distribution on $(\mathbb{R}^+, \mathcal{S})$.

$\mathcal{B}$: Borel σ-algebra on $\mathbb{R}$ or parts thereof.

Big-Oh notation : $A(x) = O(B(x))$, read "$A(x)$ is of order (or big-Oh) $B(x)$", means there exists a $C > 0$ and an x_0 such that for all $x \geq x_0$, $|A(x)| \leq CB(x)$. This is also written $A(x) \ll B(x)$ or $B(x) \gg A(x)$.

$\mathbb{C}$: the set of complex numbers: $\{z : z = x + iy,\ x, y \in \mathbb{R}\}$.

C^ℓ : the set of all ℓ times continuously differentiable functions, $\ell \in \mathbb{N}_0$.

C^∞ : the set of all smooth (i.e., infinitely differentiable) functions; $C^\infty = \bigcap_{\ell \geq 0} C^\ell$.

$D_1, D_2, D_3, \ldots$: the first, second, third, ... significant decimal digit.

$D_m^{(b)}$: the m-th significant digit base b.

$\mathbb{E}[X]$ (or $\mathbb{E}X$) : the expectation of X.

$e(x)$: $e(x) = e^{2\pi i x}$.

$f_*\mathbb{P}$: a probability measure on $\mathbb{R}$ induced by $\mathbb{P}$ and the measurable function $f : \Omega \to \mathbb{R}$, via $f_*\mathbb{P}(\cdot) := \mathbb{P}\big(f^{-1}(\cdot)\big)$.

F_n: $\{F_n\}$ is the sequence of Fibonacci numbers, $\{F_n\} = \{0, 1, 1, 2, 3, 5, 8, \ldots\}$ ($F_{n+2} = F_{n+1} + F_n$ with $F_0 = 0$ and $F_1 = 1$).

F_P, F_X : the distribution functions of P and X.

i : $i = \sqrt{-1}$.

i.i.d. : independent, identically distributed (sequence or family of random variables); often one writes i.i.d.r.v.

$\Im z$: see $\Re z$.

infimum : the infimum of a set, denoted $\inf_n x_n$, is the largest number c (if one exists) such that $x_n \geq c$ for all n, and for any $\epsilon > 0$ there is some n_0 such that $x_{n_0} < c + \epsilon$. If the sequence has finitely many terms, the infimum is the same as the minimum value.

j : in some chapters $j = \sqrt{-1}$ (this convention is frequently used in engineering).

Leb : Lebesgue measure.

Little-oh notation : $A(x) = o(B(x))$, read "$A(x)$ is little-Oh of $B(x)$", means $\lim_{x\to\infty} A(x)/B(x) = 0$.

$L^1(\mathbb{R})$: all $f : \mathbb{R} \longrightarrow \mathbb{C}$ which are measurable and Lebesgue integrable.

$\log$: usually the natural logarithm, though in some chapters it is the logarithm base 10.

$\ln$: the natural logarithm.

$\mathbb{N}$: the set of natural numbers: $\{0, 1, 2, 3, \dots\}$.

$\mathbb{N}_0$: the set of positive natural number: $\{1, 2, 3, \dots\}$.

N_f : the Newton map associated with a differentiable function f.

$o(\cdot)$, $O(\cdot)$: see 'Little-oh' and 'Big-Oh' notation, respectively.

$O_T(x_0)$: the orbit of x_0 under the map T, possibly nonautonomous.

$\{p_n\}$: the set of prime numbers: 2, 3, 5, 7, 11, 13, ...

P : probability measure on $(\mathbb{R}, \mathcal{B})$, possibly random.

P_X : the distribution of the random variable X.

Prob (or Pr) : a probability function on a probability space.

$\mathbb{Q}$: the set of rational numbers: $\{x : x = \frac{p}{q}, p, q \in \mathbb{Z}, q \neq 0\}$.

$\mathbb{R}$: the set of real numbers.

$\mathbb{R}^+$: the set of positive real numbers.

$\Re z$, $\Im z$: the real and imaginary parts of $z \in \mathbb{C}$; if $z = x + iy$, $\Re z = x$ and $\Im z = y$.

S : the significand function: if $x > 0$ then $x = S(x) \cdot 10^{k(x)}$, where $S(x) \in [1, 10)$ and $k(x) \in \mathbb{Z}$; more generally one can study the significand function S_B in base B.

$\mathcal{S}$: the significand σ-algebra.

supremum : given a sequence $\{x_n\}_{n=1}^{\infty}$, the supremum of the set, denoted $\sup_n x_n$, is the smallest number c (if one exists) such that $x_n \leq c$ for all n, and for any $\epsilon > 0$ there is some n_0 such that $x_{n_0} > c - \epsilon$. If the sequence has finitely many terms, the supremum is the same as the maximum value.

u.d. mod 1 : uniformly distributed modulo one.

$\mathrm{Var}(X)$ (or $\mathrm{var}(X)$) : the variance of the random variable X, assuming the expected value of X is finite; $\mathrm{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2]$.

$\mathbb{W}$: the set of whole numbers: $\{1, 2, 3, 4, \dots\}$.

$X_n \stackrel{\mathcal{D}}{\to} X$: (X_n) converges in distribution to X.

$X_n \stackrel{a.s.}{\to} X$: (X_n) converges to X almost surely.

$\overline{z}$, $|z|$: the conjugate and absolute value of $z \in \mathbb{C}$.

$\mathbb{Z}$: the set of integers: $\{\dots, -2, -1, 0, 1, 2, \dots\}$.

$\mathbb{Z}^+$: the set of non-negative integers, $\{0, 1, 2, \dots\}$.

PART 1

General Theory I: Basis of Benford’s Law

Chapter One

A Quick Introduction to Benford's Law

Steven J. Miller[1]

A couple of important points.

- There are many problems that would fit in multiple chapters. To help both the instructors and the readers, we have decided to collect them here. Thus, some of the exercises in this chapter will be far more accessible after reading later parts of the book.

- In Mathematica, if you define the following function you can then use it to find the first digit:

```
firstdigit[x_] := Floor[10^Mod[Log[10,x],1]]
```

  (a similar function is definable in other languages, but the syntax will differ slightly).

Exercise 1.1. *If X is Benford base 10, find the probability that its significand starts 2.789.*

Exercise 1.2. *If X is Benford base 10, find the probability that its significand starts with 7.5 (in other words, its significand is in $[7.5, 7.6)$).*

Exercise 1.3. *If X is Benford base 10, find the probability that its significant has no 7's in the first k digits (thus a significand of 1.701 would have no 7 in its first digit, but it would have a 7 in its first two digits.*

Exercise 1.4. *Consider α^n for various α and various ranges of n; for example, take $\alpha \in \{2, 3, 5, 10, \sqrt{2}, \sqrt{5}, \sqrt{10}, \pi, e, \gamma\}$ (here γ is the Euler-Mascheroni constant, see* `http://en.wikipedia.org/wiki/Euler-Mascheroni_constant` *for a description and properties), and let n go from 1 to N, where $N \in \{10^3, 10^5, 10^7\}$. Which of these data sets do you expect to be Benford? Why or why not. Read up about chi-square goodness of fit tests (see for example* `http://en.wikipedia.org/wiki/Pearson_chi_square`*) and compare the observed frequencies with the Benford probabilities.*

---

[1]Department of Mathematics and Statistics, Williams College, Williamstown, MA 01267. The author was partially supported by NSF grants DMS0970067 and DMS1265673.

Exercise 1.5. *Revisit the previous problem with more values of N. The problem is there we looked at three snapshots of the behavior; it is far more interesting to plot the chi-square values as a function of N, for N ranging from say 100 to 10^7 or more. You will see especially interesting behavior if you look at the first digits of π^n.*

Exercise 1.6. *We have seen that Benford behavior of a sequence is related to equidistribution of its logarithm. Thus, in the previous problem it may be useful to look at a log-log plot. Thus instead of plotting the chi-square value against the upper bound N, plot the logarithm of the chi-square value against* $\log N$.

Exercise 1.7. *Frequently taking logarithms helps illuminate relationships. For example, Kepler's third law (see* `http://www.physicsclassroom.com/class/circles/Lesson-4/Kepler-s-Three-Laws`*) says that the square of the time it takes a planet to orbit a sun is proportional to the cube of the semi-major axis. Find data for these quantities for the eight planets in our system (or nine if you count Pluto!) and plot them, and then do a log-log plot. A huge advantage of log-log plots is that linear relations are easy to observe and estimate; try and find the best fit line here, and note that the slope of the line should be close to 1.5 (if T is the period and L is the length of the semi-major axis, Kepler's third law is that there is a constant C such that $T^2 = CL^3$, or equivalently $T = CL^{3/2}$, or $\log T = \frac{3}{2}\log L + \log C$). Revisit the original plot, and try to see that it supports T^2 is proportional to L^3!*

Exercise 1.8. *Prove the log-laws: If $\log_b x_i = y_i$ and $r > 0$ then*

- $\log_b b = 1$ *and* $\log_b 1 = 0$ *(note $\log_b x = y$ means $x = b^y$);*
- $\log_b(x^r) = r \log_b x$;
- $\log_b(x_1 x_2) = \log_b x_1 + \log_b x_2$ *(the logarithm of a product is the sum of the logarithms);*
- $\log_b(x_1/x_2) = \log_b x_1 - \log_b x_2$ *(the logarithm of a quotient is the difference of the logarithms; this follows directly from the previous two log-laws);*
- $\log_c x = \log_b x / \log_b c$ *(this is the change of base formula).*

Exercise 1.9. *The last log-law (the change of base formula) is often forgotten, but is especially important. It tells us that if we can compute logarithms in one base then we can compute them in any base. In other words, it suffices to create just* one *table of logarithms, so we only need to find one base where we can easily compute logarithms. What base do you think that is, and how would you compute logarithms of arbitrary positive real numbers?*

Exercise 1.10. *The previous problem is similar to issues that arise in probability textbooks. These books only provide tables of probabilities of random variables drawn from a normal distribution*[2]*, as one can convert from such a table to probabilities for any other random variable. One such table is online here:*

[2] The random variable X is normally distributed with mean μ and variance σ^2 if its probability density function is $f(x;\mu,\sigma) = \exp\left(-(x-\mu)^2/(2\sigma^2)\right)/\sqrt{2\pi\sigma^2}$.

http://www.mathsisfun.com/data/standard-normal-distribution-table.html. Use a standard table to determine the probability a normal random variable with mean $\mu = 5$ *and variance* $\sigma^2 = 16$ *(so the standard deviation is* $\sigma = 4$*) takes on a value between -3 and 7. Thus, similar to the change of base formula, there is an enormous computational savings as we only need to compute probabilities for one normal distribution.*

Exercise 1.11. *Prove* $\frac{d}{dx} \log_b x = \frac{1}{x \log b}$. Hint: first do this when $b = e$, the base of the natural logarithms; use $e^{\log x} = x$ and the chain rule.

Exercise 1.12. *Revisit the first two problems, but now consider some other sequences, such as* $n!$, $\cos(n)$ *(in radians of course as otherwise the sequence is periodic),* n^2, n^3, $n^{\log n}$, $n^{\log \log n}$, $n^{\log \log \log n}$, n^n. In some situations $\log_4$ does not mean the logarithm base 4, but rather four iterations of the logarithm function. It might be interesting to investigating $n^{\log_{f(n)} n}$ under this definition for various integer-valued functions f.

Exercise 1.13. *Revisit the previous problem but for some recurrence relations. For example, try the Fibonacci numbers (*$F_{n+2} = F_{n+1} + F_n$ *with* $F_0 = 0$ *and* $F_1 = 1$*) and some other relations, such as:*

- *Catalan numbers:* $C_n = \frac{1}{n+1}\binom{2n}{n}$*; these satisfy a more involved recurrence (see http://en.wikipedia.org/wiki/Catalan_number).*
- *Squaring Fibonaccis:* $G_{n+2} = G_{n+1}^2 + G_n^2$ *with* $G_0 = 0$ *and* $G_1 = 1$.
- F_p *where* p *is a prime (i.e., only look at the Fibonaccis at a prime index).*
- *The logistic map:* $x_{n+1} = r x_n (1 - x_n)$ *for various choices of* r *and various starting values* x_0 *(see http://en.wikipedia.org/wiki/Recurrence_relation).*
- *Newton's method for the* difference *between the* n^{th} *prediction and the true value. For example, to find the square-root of* α *one uses* $x_{n+1} = \frac{1}{2}\left(x_n + \frac{\alpha}{x_n}\right)$*, and thus we would study the distribution of leading digits of* $|\sqrt{\alpha} - x_n|$*. One could also look at other roots, other numbers, or more complicated functions. For more on Newton's method, see http://mathworld.wolfram.com/NewtonsMethod.html.*
- *The* $3x + 1$ *Map:* $x_{n+1} = 3x_n + 1$ *if* x_n *is odd and* $x_n/2$ *if* x_n *is even (though some authors use a slightly different definition, where for* x_n *even one instead lets* $x_{n+1} = x_n/2^d$*, where* d *is the highest power of 2 dividing* x_n*). It is conjectured that no matter what positive starting seed* x_0 *you take, eventually* x_n *cycles among 4, 2 and 1 for* n *sufficiently large (or is identically 1 from some point onward if we use the second definition). We return to this problem in Chapter 3.*

For the remaining problems, whenever a data set satisfies Benford's law we mean the *strong* version of the law. This means the cumulative distribution function of the significand is $F_X(s) = \log_{10}(s)$ for $s \in [1, 10)$, which implies that the probability of a first digit of d is $\log_{10}(1 + 1/d)$.

Exercise 1.14. *If a data set satisfies (the strong version of) Benford's law base 10, what are the probabilities of all pairs of leading digits? In other words, what is the probability the first two digits are* $d_1 d_2$ *(in that order)? What if instead our set were Benford base* b*?*

Exercise 1.15. *Let X be a random variable that satisfies (the strong version of) Benford's law. What is the probability that the second digit is d? Note here that the possible values of d range from 0 to 9.*

Exercise 1.16. *Building on the previous problem, compute the probability that a random variable satisfying the strong version of Benford's law has its k^{th} digit equal to d. If we denote these probabilities by $p_k(d)$, what is $\lim_{k\to\infty} p_k(d)$? Prove your claim.*

Exercise 1.17. *Find a data set that is spread over several orders of magnitude, and investigate its Benfordness. For example, stock prices or volume traded on a company that has been around for decades.*

Exercise 1.18. *Look at some of the data sets from the previous exercises that were* not *Benford, and see what happens if you multiply them together. For example, consider $n^2 \cdot \cos(n)$ (in radians), or $n^2 \sqrt{10}^n \cos(n)$, or even larger products. Does this support the claim in the chapter that products of random variables tend to converge to Benford behavior?*

Exercise 1.19. *Let $\mu_{k;b}$ denote the mean of significands of k digits of random variables perfectly satisfying Benford's law, and let μ_b denote the mean of the significands of random variables perfectly following Benford's law. What is $\mu_{k;b}$ for $k \in 1, 2, 3$? Does $\mu_{k;b}$ converge to μ_b? If yes, bound $|\mu_{k;b} - \mu_b|$ as a function of k.*

Exercise 1.20. *Benford's law can be viewed as the distribution on significands arising from the density $p(x) = \frac{1}{x\log(10)}$ on $[1, 10)$ (and 0 otherwise). More generally, consider densities $p_r(x) = C_r/x^r$ for $x \in [1, 10)$ and 0 otherwise with $r \in (-\infty, \infty)$, where C_r is a normalization constant so that the density integrates to 1. For each r, calculate the probability of observing a first digit of d, and calculate the expected value of the first digit.*

Chapter Three

Fourier Analysis and Benford's Law

Steven J. Miller[1]

3.1 PROBLEMS FROM INTRODUCTION TO FOURIER ANALYSIS

The following exercises are from the chapter "An Introduction to Fourier Analysis", from the book *An Invitation to Modern Number Theory* (Princeton University Press, Steven J. Miller and Ramin Takloo-Bighash). This chapter is available online on the webpage for this book (go to the links for Chapter 3).

Exercise 3.1. *Prove e^x converges for all $x \in \mathbb{R}$ (even better, for all $x \in \mathbb{C}$). Show the series for e^x also equals*

$$\lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n, \tag{3.1}$$

which you may remember from compound interest problems.

Exercise 3.2. *Prove, using the series definition, that $e^{x+y} = e^x e^y$ and calculate the derivative of e^x.*

Exercise 3.3. *Let f, g and h be continuous functions on $[0,1]$, and $a, b \in \mathbb{C}$. Prove*

1. *$\langle f, f\rangle \geq 0$, and equals 0 if and only if f is identically zero;*
2. *$\langle f, g\rangle = \overline{\langle g, f\rangle}$;*
3. *$\langle af + bg, h\rangle = a\langle f, h\rangle + b\langle g, h\rangle$.*

Exercise 3.4. *Find a vector $\vec{v} = \binom{v_1}{v_2} \in \mathbb{C}^2$ such that $v_1^2 + v_2^2 = 0$, but $\langle \vec{v}, \vec{v}\rangle \neq 0$.*

Exercise 3.5. *Prove x^n and x^m are not perpendicular on $[0,1]$. Find a $c \in \mathbb{R}$ such that $x^n - cx^m$ is perpendicular to x^m; c is related to the projection of x^n in the direction of x^m.*

Exercise 3.6 (Important). *Show for $m, n \in \mathbb{Z}$ that*

$$\langle e_m(x), e_n(x)\rangle \;=\; \begin{cases} 1 & \text{if } m = n \\ 0 & \text{otherwise.} \end{cases} \tag{3.2}$$

---

[1]Department of Mathematics and Statistics, Williams College, Williamstown, MA 01267. The author was partially supported by NSF grants DMS0970067 and DMS1265673.

Exercise 3.7. *Let f and g be periodic functions with period a. Prove $\alpha f(x) + \beta g(x)$ is periodic with period a.*

Exercise 3.8. *Prove any function can be written as the sum of an even and an odd function.*

Exercise 3.9. *Show*

$$\langle f(x) - \widehat{f}(n)e_n(x), e_n(x)\rangle \;=\; 0. \tag{3.3}$$

This agrees with our intuition: after removing the projection in a certain direction, what is left is perpendicular to that direction.

Exercise 3.10. *Prove*

1. $\langle f(x) - S_N(x), e_n(x)\rangle = 0$ *if* $|n| \le N$.
2. $|\widehat{f}(n)| \le \int_0^1 |f(x)|dx$.
3. *Bessel's Inequality: if* $\langle f, f\rangle < \infty$ *then* $\sum_{n=-\infty}^{\infty} |\widehat{f}(n)|^2 \le \langle f, f\rangle$.
4. *Riemann-Lebesgue Lemma: if* $\langle f, f\rangle < \infty$ *then* $\lim_{|n|\to\infty} \widehat{f}(n) = 0$ *(this holds for more general f; it suffices that* $\int_0^1 |f(x)|dx < \infty$*).*
5. *Assume f is differentiable k times; integrating by parts, show* $|\widehat{f}(n)| \ll \frac{1}{n^k}$ *and the constant depends only on f and its first k derivatives.*

Exercise 3.11. *Let $h(x) = f(x)+g(x)$. Does $\widehat{h}(n) = \widehat{f}(n)+\widehat{g}(n)$? Let $k(x) = f(x)g(x)$. Does $\widehat{k}(n) = \widehat{f}(n)\widehat{g}(n)$?*

Exercise 3.12. *Remark 11.2.4 shows that if $\langle f, f\rangle, \langle g, g\rangle < \infty$ then the dot product of f and g exists: $\langle f, g\rangle < \infty$. Do there exist $f, g : [0, 1] \to \mathbb{C}$ such that $\int_0^1 |f(x)|dx, \int_0^1 |g(x)|dx < \infty$ but $\int_0^1 f(x)\overline{g}(x)dx = \infty$? Is $f \in L^2([0, 1])$ a stronger or an equivalent assumption as $f \in L^1([0, 1])$?*

Exercise 3.13. *Define*

$$A_N(x) \;=\; \begin{cases} N & \text{for } |x| \le \frac{1}{N} \\ 0 & \text{otherwise.} \end{cases} \tag{3.4}$$

Prove A_N is an approximation to the identity on $[-\frac{1}{2}, \frac{1}{2}]$. If f is continuously differentiable and periodic with period 1, calculate

$$\lim_{N\to\infty} \int_{-\frac{1}{2}}^{\frac{1}{2}} f(x)A_N(x)dx. \tag{3.5}$$

Exercise 3.14. *Let $A(x)$ be a non-negative function with $\int_{\mathbb{R}} A(x)dx = 1$. Prove $A_N(x) = N \cdot A(Nx)$ is an approximation to the identity on $\mathbb{R}$.*

Exercise 3.15 (Important)**.** *Let $A_N(x)$ be an approximation to the identity on $[-\frac{1}{2}, \frac{1}{2}]$. Let $f(x)$ be a continuous function on $[-\frac{1}{2}, \frac{1}{2}]$. Prove*

$$\lim_{N\to\infty} \int_{-\frac{1}{2}}^{\frac{1}{2}} f(x)A_N(x)dx \;=\; f(0). \tag{3.6}$$

Exercise 3.16. *Prove the two formulas above. The geometric series formula will be helpful:*

$$\sum_{n=N}^{M} r^n \;=\; \frac{r^N - r^{M+1}}{1-r}. \tag{3.7}$$

Exercise 3.17. *Show that the Dirichlet kernels are* not *an approximation to the identity. How large are* $\int_0^1 |D_N(x)|dx$ *and* $\int_0^1 D_N(x)^2 dx$*?*

Exercise 3.18. *Prove the Weierstrass Approximation Theorem implies the original version of Weierstrass' Theorem.*

Exercise 3.19. *Let* $f(x)$ *be periodic function with period* 1*. Show*

$$S_N(x_0) \;=\; \int_{-\frac{1}{2}}^{\frac{1}{2}} f(x) D_N(x-x_0)dx \;=\; \int_{-\frac{1}{2}}^{\frac{1}{2}} f(x_0-x) D_N(x)dx. \tag{3.8}$$

Exercise 3.20. *Let* $\widehat{f}(n) = \frac{1}{2^{|n|}}$*. Does* $\sum_{-\infty}^{\infty} \widehat{f}(n) e_n(x)$ *converge to a continuous, differentiable function? If so, is there a simple expression for that function?*

Exercise 3.21. *Fill in the details for the above proof. Prove the result for all* f *satisfying* $\int_0^1 |f(x)|^2 dx < \infty$*.*

Exercise 3.22. *If* $\int_0^1 |f(x)|^2 dx < \infty$*, show Bessel's Inequality implies there exists a* B *such that* $|\widehat{f}(n)| \le B$ *for all* n*.*

Exercise 3.23. *Though we used* $|a+b|^2 \le 4|a|^2 + 4|b|^2$*, any bound of the form* $c|a|^2 + c|b|^2$ *would suffice. What is the smallest* c *that works for all* $a, b \in \mathbb{C}$*?*

Exercise 3.24. *Let* $f(x) = \frac{1}{2} - |x|$ *on* $[-\frac{1}{2}, \frac{1}{2}]$*. Calculate* $\sum_{n=0}^{\infty} \frac{1}{(2n+1)^2}$*. Use this to deduce the value of* $\sum_{n=1}^{\infty} \frac{1}{n^2}$*. This is often denoted* $\zeta(2)$ *(see Exercise 3.1.7). See [BP][2] for connections with continued fractions, and [Kar][3] for connections with quadratic reciprocity.*

Exercise 3.25. *Let* $f(x) = x$ *on* $[0, 1]$*. Evaluate* $\sum_{n=1}^{\infty} \frac{1}{n^2}$*.*

Exercise 3.26. *Let* $f(x) = x$ *on* $[-\frac{1}{2}, \frac{1}{2}]$*. Prove* $\frac{\pi}{4} = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{(2n-1)^2}$*. See also Exercise 3.3.29; see Chapter* 11 *of [BB][4] or [Sc][5] for a history of calculations of* π*.*

Exercise 3.27. *Find a function to determine* $\sum_{n=1}^{\infty} \frac{1}{n^4}$*.*

---

[2] E. Bombieri and A. van der Poorten, *Continued fractions of algebraic numbers*. Pages 137–152 in *Computational Algebra and Number Theory (Sydney, 1992)*, Mathematical Applications, Vol. 325, Kluwer Academic, Dordrecht, 1995.

[3] A. Karlsson, *Applications of heat kernels on Abelian groups:* $\zeta(2n)$*, quadratic reciprocity, Bessel integral*, preprint.

[4] J. Borwein and P. Borwein, *Pi and the AGM: A Study in Analytic Number Theory and Computational Complexity*, John Wiley and Sons, New York, 1987.

[5] P. Schumer, *Mathematical Journeys*, Wiley-Interscience, John Wiley & Sons, New York, 2004.

Exercise 3.28. *Show the Gaussian* $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-(x-\mu)^2/2\sigma^2}$ *is in* $\mathcal{S}(\mathbb{R})$ *for any* $\mu, \sigma \in \mathbb{R}$.

Exercise 3.29. *Let* $f(x)$ *be a Schwartz function with compact support contained in* $[-\sigma, \sigma]$ *and denote its Fourier transform by* $\widehat{f}(y)$. *Prove for any integer* $A > 0$ *that* $|\widehat{f}(y)| \leq c_f y^{-A}$, *where the constant* c_f *depends only on* f, *its derivatives and* σ. *As such a bound is useless at* $y = 0$, *one often derives bounds of the form* $|\widehat{f}(y)| \leq \frac{\widetilde{c_f}}{(1+|y|)^A}$.

Exercise 3.30. *Consider*

$$f(x) \;=\; \begin{cases} n^6\left(\frac{1}{n^4} - |n-x|\right) & \text{if } |x-n| \leq \frac{1}{n^4} \text{ for some } n \in \mathbb{Z} \\ 0 & \text{otherwise.} \end{cases} \tag{3.9}$$

Show $f(x)$ *is continuous but* $F(0)$ *is undefined. Show* $F(x)$ *converges and is well defined for any* $x \notin \mathbb{Z}$.

Exercise 3.31. *Prove Lemma 11.4.8.*

Exercise 3.32. *For what weaker assumptions on* f, f', f'' *is the conclusion of Lemma 11.4.10 still true?*

Exercise 3.33. *One cannot always interchange orders of integration. For simplicity, we give a sequence* a_{mn} *such that* $\sum_m(\sum_n a_{m,n}) \neq \sum_n(\sum_m a_{m,n})$. *For* $m, n \geq 0$ *let*

$$a_{m,n} \;=\; \begin{cases} 1 & \text{if } n = m \\ -1 & \text{if } n = m+1 \\ 0 & \text{otherwise.} \end{cases} \tag{3.10}$$

Show that the two different orders of summation yield different answers (the reason for this is that the sum of the absolute value of the terms diverges).

Exercise 3.34. *The example in the previous remark has* $\lim_{n\to\infty} \max_x |f_n(x)| = \infty$; *in other words, there is no* M *such that* $|f_n(x)| \leq M$ *for all* M *and* x. *Find a family of functions* $f_n(x)$ *such that*

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} f_n(x)dx \;\neq\; \int_{-\infty}^{\infty} \lim_{n\to\infty} f_n(x)dx \tag{3.11}$$

and each $f_n(x)$ *and* $f(x)$ *is continuous and* $|f_n(x)|, |f(x)| \leq M$ *for some* M *and all* x.

Exercise 3.35. *Let* f, g *be continuous functions on* $I = [0,1]$ *or* $I = \mathbb{R}$. *Show if* $\langle f, f\rangle, \langle g, g\rangle < \infty$ *then* $h = f * g$ *exists.* Hint: *Use the Cauchy-Schwarz inequality. Show further that* $\widehat{h}(n) = \widehat{f}(n)\widehat{g}(n)$ *if* $I = [0,1]$ *or if* $I = \mathbb{R}$. *Thus the Fourier transform converts convolution to multiplication.*

Exercise 3.36. *Prove (11.77).*

Exercise 3.37 (Important)**.** *If for all* $i = 1, 2, \ldots$ *we have* $\langle f_i, f_i\rangle < \infty$, *prove for all* i *and* j *that* $\langle f_i * f_j, f_i * f_j\rangle < \infty$. *What about* $f_1 * (f_2 * f_3)$ *(and so on)? Prove* $f_1 * (f_2 * f_3) = (f_1 * f_2) * f_3$. *Therefore convolution is associative, and we may write* $f_1 * \cdots * f_N$ *for the convolution of* N *functions.*

Exercise 3.38. *Suppose $X_1, \dots, X_N$ are i.i.d.r.v. from a probability distribution p on $\mathbb{R}$. Determine the probability that $X_1 + \cdots + X_N \in [a, b]$. What must be assumed about p for the integrals to converge?*

Exercise 3.39. *One useful property of the Fourier transform is that the derivative of $\widehat{g}$ is the Fourier transform of $2\pi i x g(x)$; thus, differentiation (hard) is converted to multiplication (easy). Explicitly, show*

$$\widehat{g}'(y) \ = \ \int_{-\infty}^{\infty} 2\pi i x \cdot g(x) e^{-2\pi i x y} dx. \tag{3.12}$$

If g is a probability density, note $\widehat{g}'(0) = -2\pi i \mathbb{E}[x]$ and $\widehat{g}''(0) = -4\pi^2 \mathbb{E}[x^2]$.

Exercise 3.40. *If $B(x) = A(cx)$ for some fixed $c \neq 0$, show $\widehat{B}(y) = \frac{1}{c}\widehat{A}\left(\frac{y}{c}\right)$.*

Exercise 3.41. *Show that if the probability density of $X_1 + \cdots + X_N = x$ is $(p * \cdots * p)(x)$ (i.e., the distribution of the sum is given by $p * \cdots * p$), then the probability density of $\frac{X_1 + \cdots + X_N}{\sqrt{N}} = x$ is $(\sqrt{N}p * \cdots * \sqrt{N}p)(x\sqrt{N})$. By Exercise 3.40, show*

$$FT\left[(\sqrt{N}p * \cdots * \sqrt{N}p)(x\sqrt{N})\right](y) \ = \ \left[\widehat{p}\left(\frac{y}{\sqrt{N}}\right)\right]^N. \tag{3.13}$$

Exercise 3.42. *Show for any fixed y that*

$$\lim_{N\to\infty}\left[1 - \frac{2\pi^2 y^2}{N} + O\left(\frac{y^3}{N^{3/2}}\right)\right]^N \ = \ e^{-2\pi^2 y^2}. \tag{3.14}$$

Exercise 3.43. *Show that the Fourier transform of $e^{-2\pi^2 y^2}$ at x is $\frac{1}{\sqrt{2\pi}}\, e^{-x^2/2}$.* Hint: *This problem requires contour integration from complex analysis.*

Exercise 3.44. *Modify the proof to deal with the case of p having mean μ and variance σ^2.*

Exercise 3.45. *For reasonable assumptions on p, estimate the rate of convergence to the Gaussian.*

Exercise 3.46. *Let p_1, p_2 be two probability densities satisfying (11.79). Consider $S_N = X_1 + \cdots + X_N$, where for each i, X_1 is equally likely to be drawn randomly from p_1* or *p_2. Show the Central Limit Theorem is still true in this case. What if we instead had a fixed, finite number of such distributions $p_1, \dots, p_k$, and for each i we draw X_i from p_j with probability q_j (of course, $q_1 + \cdots + q_k = 1$)?*

Exercise 3.47 (Gibbs Phenomenon)**.** *Define a periodic with period 1 function by*

$$f(x) \ = \ \begin{cases} -1 & \text{if } -\frac{1}{2} \le x < 0 \\ 1 & \text{if } 0 \le x < \frac{1}{2}. \end{cases} \tag{3.15}$$

Prove that the Fourier coefficients are

$$\widehat{f}(n) \ = \ \begin{cases} 0 & \text{if } n \text{ is even} \\ \frac{4}{n\pi i} & \text{if } n \text{ is odd.} \end{cases} \tag{3.16}$$

Show that the N^{th} partial Fourier series $S_N(x)$ converges pointwise to $f(x)$ wherever f is continuous, but overshoots and undershoots for x near 0. Hint: *Express the series expansion for $S_N(x)$ as a sum of sines. Note $\frac{\sin(2m\pi x)}{2m\pi} = \int_0^x \cos(2m\pi t)dt$. Express this as the real part of a geometric series of complex exponentials, and use the geometric series formula. This will lead to*

$$S_{2N-1}(x) \;=\; 8\int_0^x \Re\left(\frac{1}{2i}\frac{e^{4n\pi it}-1}{\sin(2\pi t)}\right)dt \;=\; 4\int_0^x \frac{\sin(4n\pi t)}{\sin(2\pi t)}dt, \tag{3.17}$$

which is about 1.179 (or an overshoot of about 18%) when $x = \frac{1}{4n\pi}$. What can you say about the Fejér series $T_N(x)$ for x near 0?

Exercise 3.48 (Nowhere Differentiable Function)**.** *Weierstrass constructed a continuous but nowhere differentiable function! We give a modified example and sketch the proof. Consider*

$$f(x) \;=\; \sum_{n=0}^{\infty} a^n \cos(2^n \cdot 2\pi x), \quad \frac{1}{2} < a < 1. \tag{3.18}$$

Show f is continuous but nowhere differentiable. Hint: *First show $|a| < 1$ implies f is continuous. Our claim on f follows from: if a periodic continuous function g is differentiable at x_0 and $\widehat{g}(n) = 0$ unless $n = \pm 2^m$, then there exists C such that for all n, $|\widehat{g}(n)| \le Cn2^{-n}$. To see this, show it suffices to consider $x_0 = 0$ and $g(0) = 0$. Our assumptions imply that $(g, e_m) = 0$ if $2^{n-1} < m < 2^{n+1}$ and $m \ne 2^n$. We have $\widehat{g}(2^n) = (g, e_{2^n}F_{2^{n-1}}(x))$ where F_N is the Fejér kernel. The claim follows from bounding the integral $(g, e_{2^n}F_{2^{n-1}}(x))$. In fact, more is true: Baire showed that, in a certain sense, "most" continuous functions are nowhere differentiable! See, for example, [Fol][6].*

Exercise 3.49 (Isoperimetric Inequality)**.** *Let $\gamma(t) = (x(t), y(t))$ be a smooth closed curve in the plane; we may assume it is parametrized by arc length and has length 1. Prove the enclosed area A is largest when $\gamma(t)$ is a circle.* Hint: *By Green's Theorem (Theorem A.2.9),*

$$\oint_\gamma x dy - y dx \;=\; 2\text{Area}(A). \tag{3.19}$$

The assumptions on $\gamma(t)$ imply $x(t), y(t)$ are periodic functions with Fourier series expansions and $\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 = 1$. Integrate this equality from $t = 0$ to $t = 1$ to obtain a relation among the Fourier coefficients of $\frac{dx}{dt}$ and $\frac{dx}{dt}$ (which are related to those of $x(t)$ and $y(t)$); (3.19) *gives another relation among the Fourier coefficients. These relations imply $4\pi\text{Area}(A) \le 1$ with strict inequality unless the Fourier coefficients vanish for $|n| > 1$. After some algebra, one finds this implies we have a strict inequality unless γ is a circle.*

Exercise 3.50 (Applications to Differential Equations)**.** *One reason for the introduction of Fourier series was to solve differential equations. Consider the vibrating string problem: a unit string with endpoints fixed is stretched into some initial position and then released;*

[6] G. Folland, *Real Analysis: Modern Techniques and Their Applications*, 2nd edition, Pure and Applied Mathematics, Wiley-Interscience, New York, 1999.

describe its motion as time passes. Let $u(x,t)$ denote the vertical displacement from the rest position x units from the left endpoint at time t. For all t we have $u(0,t) = u(1,t) = 0$ as the endpoints are fixed. Ignoring gravity and friction, for small displacements Newton's laws imply

$$\frac{\partial^2 u(x,t)}{\partial x^2} \;=\; c^2 \frac{\partial^2 u(x,t)}{\partial t^2}, \tag{3.20}$$

where c depends on the tension and density of the string. Guessing a solution of the form

$$u(x,t) \;=\; \sum_{n=1}^{\infty} a_n(t) \sin(n\pi x), \tag{3.21}$$

solve for $a_n(t)$.

One can also study problems on $\mathbb{R}$ by using the Fourier Transform. Its use stems from the fact that it converts multiplication to differentiation, and vice versa: if $g(x) = f'(x)$ and $h(x) = xf(x)$, prove that $\widehat{g}(y) = 2\pi i y \widehat{f}(y)$ and $\frac{d\widehat{f}(y)}{dy} = -2\pi i \widehat{h}(y)$. This and Fourier Inversion allow us to solve problems such as the heat equation

$$\frac{\partial u(x,t)}{\partial t} \;=\; \frac{\partial^2 u(x,t)}{\partial x^2}, \quad x \in \mathbb{R},\ t > 0 \tag{3.22}$$

with initial conditions $u(x,0) = f(x)$.

3.2 PROBLEMS FROM CHAPTER 1: REVISITED

Many of the problems from Chapter 1 are appropriate here as well. In addition to re-examining those problems, consider the following.

Exercise 3.51. *Is the sequence $a_n = n^{\log n}$ Benford?*

Exercise 3.52. *In some situations $\log_4$ does not mean the logarithm base 4, but rather four iterations of the logarithm function. Investigate $n^{\log_{f(n)} n}$ under this definition for various integer-valued functions f.*

3.3 PROBLEMS FROM CHAPTER 3

Exercise 3.53. *Assume an infinite sequence of real numbers $\{x_n\}$ has its logarithms modulo 1, $\{y_n = \log_{10} x_n \bmod 1\}$, satisfying the following property: as $n \to \infty$ the proportion of y_n in any interval $[a,b] \subset [0,1]$ converges to $b - a$* **if** *$b - a > 1/2$. Prove or disprove that $\{x_n\}$ is Benford.*

Exercise 3.54. *As $\sqrt{2}$ is irrational, the sequence $\{x_n = n\sqrt{2}\}$ is uniformly distributed modulo 1. Is the sequence $\{x_n^2\}$ uniformly distributed modulo 1?*

Exercise 3.55. *Does there exist an irrational α such that α is a root of a quadratic polynomial with integer coefficients* and *the sequence $\{\alpha^n\}_{n=1}^{\infty}$ is Benford base 10?*

Exercise 3.56. *Which of the following are Benford, and why?*

- $\{nx^2\} = \{x^2, 2x^2, 3x^2, 4x^2, \dots\}$ *for* $x > 1$.
- $\{n\pi^2\} = \{\pi^2, 2\pi^2, 3\pi^2, 4\pi^2, \dots\}$.
- $\{n^2\pi^2\} = \{\pi^2, 4\pi^2, 9\pi^2, 16\pi^2, \dots\}$.
- $\{5^n + 2^n\} = \{7, 29, 133, 641, \dots\}$.
- $\{e^{1+n\pi}\} = \{e^{1+\pi}, e^{1+2\pi}, e^{1+3\pi}, e^{1+4\pi}, \dots\}$.
- $1/X$, *where* X *is a random variable uniformly distributed on* $[0, 5)$.

Exercise 3.57. *We showed a Geometric Brownian Motion is a Benford-good process; is the sum of two independent Geometric Brownian Motions Benford-good?*

The next few questions are related to a map we now describe. We showed that, suitably viewed, the $3x + 1$ Map leads to Benford behavior (or is close to Benford for almost all large starting seeds). Consider the following map. Let $R(x)$ be the number formed by writing the digits of x in reverse order. If $R(x) = x$ we say x is palindromic. If x is not a palindromic number set $P(x) = x + R(x)$, and if x is palindromic let $P(x) = x$. For a given starting seed x_0 consider the sequence where $x_{n+1} = P(x)$. It is not known if there are any x_0 such that the resulting sequence diverges to infinity, though it is believed that almost all such numbers do. The first candidate to escape is 196; for more see http://en.wikipedia.org/wiki/Lychrel_number (this process is also called “reverse-and-add”, and the candidates are called Lychrel numbers).

Exercise 3.58. *Consider the reverse-and-add map described above applied to a large starting seed. Find as good of a lower bound as you can for the number of seeds between* 10^n *and* 10^{n+1} *such that the resulting sequence stabilizes (i.e., we eventually hit a palindrome).*

Exercise 3.59. *Come up with a model to estimate the probability a given starting seed in* 10^n *and* 10^{n+1} *has its iterates under the reverse-and-add map diverge to infinity.* Hint: x plus $R(x)$ is a palindrome if and only if there are no carries when we add; thus you must estimate the probability of having no carries.

Exercise 3.60. *Investigate the Benfordness of sequences arising from the reverse-and-add map for various starting seeds. Of course the calculation is complicated by our lack of knowledge about this map, specifically we don't know even one starting seed that diverges! Look at what happens with various Lycherel numbers. For each* N *can you find a starting seed* x_0 *such that it iterates to a palindrome after* N *or more steps?*

Exercise 3.61. *Redo the previous three problems in different bases. Your answer will depend now on the base; for example, much more is known base 2 (there we can give specific starting seeds that iterate to infinity).*

Exercise 3.62. *Use the Erdös-Turan Inequality to calculate upper bounds for the discrepancy for various sequences, and use those results to prove Benford behavior. Note you need to find a sequence where you can do the resulting computation. For example, earlier we investigated* $a_n = n^{\log n}$*; are you able to do the summation for this case?*

Exercise 3.63. *Consider the analysis of products of random variables. Fix a probability* p *(maybe* $p = 1/2$*), and independent identically distributed random variables* $X_1, \dots, X_n$*. Assume as* $n \to \infty$ *the product of the* X_i*'s becomes Benford. What if now we let* $\widetilde{\mathcal{X}_n}$ *be the random variable where we toss* n *independent coins, each with probability* p*, and if the* i^{th} *toss is a head then* X_i *is in the product (if the product is empty we use the standard convention that it is then 1). Is this process Benford?*

Exercise 3.64. *Redo the previous problem, but drop the assumption that the random variables are identically distributed.*

Exercise 3.65. *Redo the previous two problems, but now allow the probability that the* i^{th} *toss is a head to depend on* i*.*

Exercise 3.66. *Consider*

$$\phi_m = \begin{cases} m & \text{if } |x - \frac{1}{8}| \le \frac{1}{2m} \\ 0 & \text{otherwise;} \end{cases} \tag{3.23}$$

this is the function from Example 3.3.5 and led to non-Benford behavior for the product. Can you write down the density for the product?

Exercise 3.67. *In the spirit of the previous problem, find other random variables where the product is not Benford.*

Exercise 3.68. *Consider a Weibull random variable where we fix the scale parameter* α *to be 1 and the translation parameter* β *to be 0; thus* $f(x;\gamma) = x^{\gamma-1}\exp(x^\gamma)$ *for* $x \ge 0$ *and is zero otherwise. Investigate the Benfordness of chaining random variables here, where the shape parameter* γ *is the output of the previous step.*

Exercise 3.69. *The methods of [JKKKM] led to good bounds for chaining exponential and uniform random variables. Can you obtain good, explicit bounds in other cases? For example, consider a binomial process with fixed parameter* p*.*

Exercise 3.70. *Apply the methods of Cuff, Lewis and Miller (for the Weibull distribution)to other random variables. Consider the generalized Gamma distribution (see*

http://en.wikipedia.org/wiki/Generalized_gamma_distribution

for more information), where the density is

$$f(x;a,d,p) = \frac{p/d^a}{\Gamma(d/p)} x^{d-1} \exp\left(-(x/a)^p\right)$$

for $x > 0$ *and 0 otherwise, where* a, d, p *are positive parameters.*

For the next few problems, let $f_r(x) - 1/(1+|x|^r)$ with $r > 1$.

Exercise 3.71. *Show that for* $r > 1$*,* $\int_{-\infty}^{\infty} f_r(x)\,dx$ *is finite. Verify that* $\int_{-\infty}^{\infty} f_r(x)\,dx = \frac{2\pi}{r}\csc\left(\frac{\pi}{r}\right)$.

Exercise 3.72. *Verify that the Fourier transform identity used in our analysis:*

$$p_r\left(e^{b+y}\right)e^{b+y} = \frac{1}{2}\sin\left(\frac{\pi}{r}\right)e^{2\pi i b y}\csc\left(\frac{\pi}{r}(1-2\pi i y)\right),$$

where $b \in [0,1]$.

PART 2
General Theory II: Distributions and Rates of Convergence

Chapter Four

Benford's Law Geometry

Lawrence Leemis[1]

Exercise 4.1. *Perform a chi-square goodness-of-fit test on the data values in Table 4.1.*

Exercise 4.2. *Let the random variable X have the Benford distribution as defined in this chapter. Find $\mathbb{E}[X]$. Next, generate one million Benford random variates and compute their sample mean. Perform this Monte Carlo experiment several times to assure that the sample means are near $\mathbb{E}[X]$.*

Exercise 4.3. *Let $T \sim$ exponential(1). Find the probability mass function of the leading digit to three-digit accuracy. Compare your results to those in Table 4.2.*

Exercise 4.4. *Redo the previous exercise, but instead of finding the probability mass function of the leading digit, find the cumulative distribution function of the significand (i.e., find the probability of observing a significand of at most s).*

Exercise 4.5. *Determine the set of conditions on a, b, and c associated with $W \sim$ triangular(a, b, c) which result in $T = 10^W$ following Benford's Law.*

Exercise 4.6. *Use R to confirm that the cumulative distribution function $F_x(x) = \mathrm{Prob}(X \leq x) = \log_{10}(x+1)$ results in a probability mass function that gives the distribution specified in Benford's Law. What is the range of x?*

Exercise 4.7. *Use R to determine if the cumulative distribution function $F_x(x) = \mathrm{Prob}(X \leq x) = x^2$ (for some range for x) results in a probability mass function that gives the distribution specified in Benford's Law. If yes, what is the range for x?*

Exercise 4.8. *Which of the following distributions of W follow Benford's Law?*

- $f_W(w) \sim U(0, 3.5)$.
- $f_W(w) \sim U(17, 117)$.
- $f_W(w) = w^3 - w^2 + w$ *for* $0 \leq w \leq 1$, *and* $1 - w^3 + w^2 - w$ *for* $1 \leq w \leq 2$.
- $f_W(w) = \sqrt{w}$ *for* $0 \leq w \leq 1$, *and* $1 - \sqrt{w-1}$ *for* $1 \leq w \leq 2$.

Exercise 4.9. *Let b_1 and b_2 be two different integers exceeding 1. Is there a probability density p on an interval I such that if a random variable X has p for its probability density function then X is Benford in both base b_1 and b_2? What if the two bases are allowed to be real numbers exceeding 1? Prove your claims.*

[1]Department of Mathematics, The College of William & Mary, Williamsburg, VA 23187.

Chapter Five

Explicit Error Bounds via Total Variation

Lutz Dümbgen and Christoph Leuenberger[1]

Exercise 5.1. *Find* $TV(\sin(x), [-\pi, \pi])$.

Exercise 5.2. *Confirm that* $TV(h, \mathbb{J}) = TV^+(h, \mathbb{J}) + TV^-(h, \mathbb{J})$.

Exercise 5.3. *Let* Y_o *and* Z *be independent random variables such that* Y_o *has a density* f_o *with* $\mathrm{TV}(f_o) < \infty$ *and* Z *has distribution* π. *Verify that* $Y := Y_o + Z$ *has density* $f(y) = \int f_o(y - z)\,\pi(dz)$ *with* $\mathrm{TV}(f) \le \mathrm{TV}(f_o)$.

Exercise 5.4. *Show that an absolutely continuous probability density* f *on* $\mathbb{R}$ *satisfies*

$$\mathrm{TV}(f)^2 \;\le\; \int \frac{f'(x)^2}{f(x)}\,dx.$$

Exercise 5.5. *Let* $\gamma_{a,\sigma}$ *be the density of the gamma distribution* $\mathrm{Gamma}(a, \sigma)$ *with shape parameter* $a > 0$ *and scale parameter* $\sigma > 0$, *i.e.,*

$$\gamma_{a,\sigma}(x) \;=\; \sigma^{-a} x^{a-1} \exp(-x/\sigma)/\Gamma(a)$$

for $x > 0$, *and* $\gamma_{a,\sigma} = 0$ *on* $(-\infty, 0]$.

(i) *Show that for* $a \ge 1$,

$$\mathrm{TV}(\gamma_{a,\sigma}) \;=\; \sigma^{-1}\,\mathrm{TV}(\gamma_{a,1}) \quad \textit{and} \quad \mathrm{TV}(\gamma_{a,1}) \;=\; 2((a-1)/e)^{a-1}/\Gamma(a).$$

(ii) *It is well-known that* $\Gamma(t+1) = (t/e)^t\sqrt{2\pi t}(1 + o(1))$ *as* $t \to \infty$ *(this is Stirling's formula). What does this imply for* $\mathrm{TV}(\gamma_{a,1})$? *Show that* $\mathrm{TV}(\gamma_{a,\sigma}) \to 0$ *as* $\sqrt{a}\,\sigma \to \infty$ *and* $a \ge 1$.

Exercise 5.6. *Let* X *be a strictly positive random variable with density* h *on* $(0, \infty)$. *Verify that* $Y := \log_B(X)$ *has density* f *given by* $f(y) = \log(B)B^y h(B^y)$ *for* $y \in \mathbb{R}$.

Exercise 5.7. *Let* X *be a random variable with distribution* $\mathrm{Gamma}(a, \sigma)$ *for some* $a, \sigma > 0$*; see Exercise 5.5*

(i) *Determine the density* $f_{a,\sigma}$ *of* $Y := \log_B(X)$. *Here you should realize that* $f_{a,\sigma}(y) = f_{a,1}(y - \log_B(\sigma))$. *Show then that*

$$\mathrm{TV}(f_{a,\sigma}) \;=\; 2\log(B)(a/e)^a/\Gamma(a).$$

What happens as $a \to \infty$?

(ii) *To understand why the leading digits of* X *are far from Benfords law for large* a, *verify that* $X = \sigma(a + \sqrt{a}Z_a)$ *for a random variable* Z_a *with mean zero and variance one. (Indeed, the density of* Z_a *converges uniformly to the standard Gaussian density as* $a \to \infty$.) *Now investigate the distribution of* $Y = \log_B(X)$ *as* $a \to \infty$.

[1]University of Bern, Bern, Switzerland and University of Fribourg, Fribourg, Switzerland respectively.

Chapter Six

Lévy Processes and BenfordŠs Law

Klaus Schürger[1]

Exercise 6.1. *Provide an example of a non-continuous cadlag function.*

Exercise 6.2. *Prove that a Weiner Process is also a Lévy Process.*

Exercise 6.3. *Prove that a Poisson Process is also a Lévy Process.*

Exercise 6.4. *Prove that the exponential Lévy Process* $\{\exp(X_t)\}$ $(t \in \mathbb{R})$ *is a martingale with respect to* $(\mathcal{F}_t) := \sigma\{X_s : s \leq t\}$ *if and only if* $\mathbb{E}[\exp(X_t)] = 1$.

Exercise 6.5. *Let* $f(t) = \mathbb{E}[\exp(it\xi)], g(t) = \mathbb{E}[\exp(it\eta)]$ $(t \in \mathbb{R})$ *be the characteristic functions of* (*real-*) *valued random variables* ξ, η $(i = \sqrt{-1})$. *Recall that* $\exp(it) = \cos t + i \sin t$ $(t \in \mathbb{R})$ *and* $\mathbb{E}[\exp(it\xi)] := \mathbb{E}[\cos(t\xi)] + i\mathbb{E}(\sin(t\xi)]$ $(t \in \mathbb{R})$. *Finally,* $\overline{a+ib} := a - ib$ $(a, b \in \mathbb{R})$ *denotes the complex conjugate of* $a + ib$. *Note that* $|f|^2(t) = f(t) \cdot \bar{f}(t)$. *Show the following.*

(a) f *is continuous,* $f(0) = 1$, *and* $|f(t)| \leq 1,\ t \in \mathbb{R}$.

(b) $\bar{f}$ *is a characteristic function.*

(c) $f \cdot g$ *is a characteristic function. Hence,* $|f|^2$*is a characteristic function.*

(d) *Let* $h_1, h_2, \ldots$ *be characteristic functions. If* $a_1 \geq 0, a_2 \geq 0, \ldots$ *are real numbers such that* $a_1 + a_2 + \cdots = 1$, *then* $a_1h_1 + a_2h_2 + \cdots$ *is a characteristic function.*

(e) *Show that every characteristic function h is non-negative definite, i.e., for all* $n \geq 2$, *real* $t_1, \ldots, t_n$ *and complex* $a_1, \ldots, a_n$ *we have that*

$$\sum_{j=1}^{n} \sum_{k=1}^{n} h(t_j - t_k) a_j \bar{a}_k \geq 0.$$

Exercise 6.6. *Show that, for each real number* $p > 0, f(z) := \cos(2\pi pz)$ $(z \in \mathbb{R})$ *is a characteristic function. Deduce that* $g(z) := (\cos(2\pi pz))^2$ $(z \in \mathbb{R})$ *is a characteristic function.*

[1]Department of Economics, University of Bonn, Adenauerallee 24-42, 53113 Bonn, Germany.

Exercise 6.7. (This exercise gives an example of a characteristic function which "wildly fluctuates".) *It follows from Exercises* 6.6 *and* 6.5(d) *that*

$$h(z) := \sum_{k=1}^{\infty} 2^{-k}(\cos(2\pi 7^k z))^2, \ z \in \mathbb{R}$$

is a characteristic function. Show that h *is of infinite total variation over each non-degenerate interval* $[a, b]$*, i.e.,*

$$\sup\left\{\sum_{k=1}^{n} |h(z_{k+1}) - h(z_k)|\right\} = \infty,$$

the supremum taken over all $n \geq 1$ *and real numbers* $a \leq z_1 < z_2 < \cdots < z_{n+1} \leq b$*.*

[*Hint: It suffices to prove the claim for intervals* $[r + 7^{-N}, r + 2 \cdot 7^{-N}]$ (*being convenient for calculations!*) *where* $N \geq 1$ *is an integer and* $r \geq 0$ *a real number. Let* $k \geq N + 1$ *and denote by* $I(k)$ *the set of integers* j *such that* $1 + (r + 7^{-N})7^k < j \leq ((r + 2 \cdot 7^{-N})7^k$*. For* $j \in I(k)$ *put* $t_{2j-1}(k) = (j - 1/4)7^{-k}, t_{2j}(k) = j \cdot 7^{-k}$*. Show, by using the inequalities* $|a + b| \geq |a| - |b|$ *and* $|(\cos b)^2 - (\cos a)^2| \leq 2|b - a|$ $(a, b \in \mathbb{R})$ *that*

$$\sum_{j \in I(k)} |h(t_{2j}(k)) - h(t_{2j-1}(k))| \geq 2(1 - \pi/5)7^{-N}(7/2)^k + \text{const.}]$$

Exercise 6.8. (*a*) *Try to guess how the integral* $\int_a^b f(z)\exp(itz)dz$ *behaves as* $t \to \infty$ *if* $f : [a, b] \to \mathbb{R}$ *is a step function of the form* $f(t) = \sum_{j=1}^{m} c_j \mathbb{1}_{[b_{j-1}, b_j)}(t)$ *where* $a \leq b_0 < b_1 < \cdots < b_m \leq b$*.*

(b) *Verify your guess when* f *is an indicator function of an interval.*

(c) *How does the above integral behave when* f *is continuous on* $[a, b]$*?*

Exercise 6.9. *Show that a Lévy measure* Q *satisfies* $Q(\mathbb{R} \setminus (-\alpha, \alpha)) < \infty$ *for all* $\alpha > 0$*.*

Exercise 6.10. *Let* X *be a Lévy process having Lévy measure* Q*. Show that, for fixed* $c > 0$ *and* $s \geq 0$*, the process* X^* *given by* $X_t^* = X_{ct+s} - X_s$ $(t \geq 0)$ *is a Lévy process having Lévy measure* $Q^* = cQ$*.*

Exercise 6.11. *Let* $N = (N_t)$ $(t \geq 0)$ *be a Poisson process with parameter* $\lambda > 0$*.*

(a) *Verify that the generating triple of* N *is given by* $(\lambda, 0, Q^*)$ *where* Q^* *has total mass* λ *concentrated on* $\{1\}$*.*

(b) *Verify* (6.15) *directly for* $X = N$*, i.e.,*

$$Q^*(A) = c^{-1}\mathbb{E}[\#\{s < t \leq s + c : \Delta N_t \in A \setminus \{0\}\}]$$

holds for all $c > 0, s \geq 0$ *and every Borel set* $A \subset \mathbb{R}$*.*

Exercise 6.12. *Let* $T_t = \sum_{j=1}^{N_t} \zeta_j$ $(t \geq 0)$ *denote the compound Poisson process of Example* 6.1.21. (*Here,* (N_t) *is a Poisson process with parameter* $\lambda > 0; \zeta_1, \zeta_2, \ldots$ *are independent random variables with a common distribution* Q_1 *such that* $Q_1(\{0\}) = 0$*. Furthermore, the processes* (ζ_n) *and* (N_t) *are independent of each other.*)

(a) *Show that the characteristic function* g_t *of* T_t $(t \geq 0)$ *is given by*

$$g_t(z) \;=\; \exp\left[\lambda t \int_{\mathbb{R}} (e^{izx} - 1) Q_1(dx)\right]$$

for all $z \in \mathbb{R}$ *and* $t \geq 0$.

(b) *It can be shown* (*see the reference in Example* 6.1.21) *that* (T_t) *is a Lévy process. Determine its generating triple* (β, σ^2, Q).

Exercise 6.13. *Let* W *be a (standard) Brownian motion* (BM). *Show that, for each* $c > 0, W^* = (cW_{t/c^2})$ *is a* BM (*scaling property*).

Exercise 6.14. *Let* $\xi \sim N(\mu, \sigma^2)$ *where* $\mu \in \mathbb{R}$ *and* $\sigma > 0$.

(a) *Deduce from* (6.26) *that the characteristic function of* ξ *is given by*

$$\mathbb{E}[\exp(iz\xi)] \;=\; \exp(i\mu z - \sigma^2 z^2/2),\; z \in \mathbb{R}.$$

(b) *Deduce from the result in* (a) *that, for all* $\mu, z \in \mathbb{R}$ *and* $\sigma > 0$,

$$\int_{-\infty}^{\infty} \cos(zx)\exp(-(x-\mu)^2/(2\sigma^2))dx \;=\; \sqrt{2\pi\sigma^2}\cos(\mu z)\exp(-\sigma^2 z^2/2)$$

and

$$\int_{-\infty}^{\infty} \sin(zx)\exp(-(x-\mu)^2/(2\sigma^2))dx \;=\; \sqrt{2\pi\sigma^2}\sin(\mu z)\exp(-\sigma^2 z^2/2).$$

Exercise 6.15. *Let* $W = (W_t)$ *be a* BM. *Put*

$$S_{t,u} \;:=\; \sup_{0 \leq s \leq u} |W_{t+s} - W_t|,\; t \geq 0, u > 0.$$

(a) *Show that* $S_{t,u}$ *is a random variable.* (*This requires a little argument since the definition of* $S_{t,u}$ *involves uncountably many random variables!*)
[*Hint: Recall that all sample paths of* W *are continuous.*]

(b) *Show that* $W_n/n \to 0$ $(n \to \infty)$ *a.s.*

(c) *Since, for each fixed* $t \geq 0, (W_{u+t} - W_t)$ $(u \geq 0)$ *is a* BM, *it follows that*

$$\text{For each } t > 0, S_{t,1} \text{ has the same distribution as } S_{0,1}. \tag{$*$}$$

Furthermore, we have that

$$P(S_{0,1} \geq a) \;\leq\; 2\exp(-a^2/2),\; a \geq 0 \tag{$**$}$$

(*see, e.g., [KaSh] or [RY]*). *Use* (*b*) *as well as* $(*)$ *and* $(**)$ *to show that*

$$W_t/t \to 0 \; (t \to \infty) \text{ a.s.}$$

[*Hint:Use the Borel-Cantelli Lemma.*]

Exercise 6.16. *Let* $\xi_1, \xi_2, \ldots$ *be independent random variables defined on some probability space* $(\Omega, \mathcal{F}, P)$, *which have a common distribution given by* $P(\xi_n = +1) = p, P(\xi_n = -1) = 1 - p =: q$ $(n \geq 1)$, *where* $0 < p < 1$. *Put* $S_n := \xi_1 + \cdots + \xi_n,\; n \geq 0$ $(S_0 = 0)$, *and let* $(\mathcal{F}_n)$ $(n \geq 0)$ *be the filtration generated by* (ξ_n). (*Note that* $\mathcal{F}_0 = \{\emptyset, \Omega\}$.)

(a) *Show that* $Y_n := (q/p)^{S_n}$ $(n \geq 0)$ *is an* $(\mathcal{F}_n)$*-martingale.*

(b) *Put* $c(\alpha) := \mathbb{E}[\exp(\alpha\xi_1)] = p\exp(\alpha) + q\exp(-\alpha)$ $(\alpha \in \mathbb{R})$. *Show that, for every fixed* $\alpha \in \mathbb{R}$,

$$Z_n \;:=\; \exp(\alpha S_n)/(c(\alpha))^n \;(n \geq 0)$$

is an $(\mathcal{F}_n)$*-martingale.*

Exercise 6.17. *Let* $\xi_1, \xi_2, \ldots$ *be independent random variables defined on the same probability space, which have a common distribution given by* $P(\xi_n = +1) = P(\xi_n = -1) = 1/2$. *Put* $S_0 = 0$ *and* $S_n = \xi_1 + \cdots + \xi_n$ $(n \geq 1)$ *which means that* (S_n) *is a simple symmetric random walk on* $\mathbb{Z}$, *starting at* 0. *Let* $(\mathcal{F}_n)$ *be the filtration generated by* (ξ_n). *Show that following two sequences are* $(\mathcal{F}_n)$*-martingales:*

(a) $(S_n^3 - 3nS_n)$.

(b) $(S_n^4 - 6nS_n^2 + 3n^2 + 2n)$.

[*Hint: Note that* $\mathbb{E}[\xi_n|\mathcal{F}_{n-1}] = \mathbb{E}[\xi_n] = 0$ *a.s.* (*since* ξ_n *is independent of* $\mathcal{F}_{n-1}$), *and that* $\mathbb{E}[S_{n-1}^2\xi_n|\mathcal{F}_{n-1}] = S_{n-1}^2\mathbb{E}[\xi_n] = 0$ *a.s.* (*since* S_{n-1} *is* $\mathcal{F}_{n-1}$*-measurable*). *Note that* $S_n = S_{n-1} + \xi_n$.]

Exercise 6.18. *Let* $(\Omega, \mathcal{F}, P)$ *be a probability space and let* $(\mathcal{F}_n)$ $(n \geq 0)$ *be any filtration on* $(\Omega, \mathcal{F})$. *In the sequel let* $Z = (Z_n)$ $(n \geq 0)$ *and* $H = (H_n)$ $(n \geq 1)$ *be sequences of random variables defined on* $(\Omega, \mathcal{F})$ *such that* Z *is adapted and* H *is predictable which means that, for all* $n \geq 1$, H_n *is* $\mathcal{F}_{n-1}$*-measurable. The sequence* $H \bullet Z$ *given by*

$$(H \bullet Z)_n \;:=\; \sum_{j=1}^{n} H_j(Z_j - Z_{j-1}),\; n \geq 0 \;((H \bullet Z)_0 = 0)$$

is called the H*-transform of* Z *or the* (*discrete*) *stochastic integral of* H *with respect to* Z. *Now let* Z *be an* $(\mathcal{F}_n)$*-martingale and assume that* $H_j(Z_j - Z_{j-1}) \in L^1,\ j = 1, 2, \ldots$ *Show that* $H \bullet Z$ *is an* $(\mathcal{F}_n)$*-martingale.*
[*Hint: Use the iteration property of conditional expectations* (*see Example* 6.1.29).]

Exercise 6.19. *Let* $W = (W_t)$ *be a* BM *and let* $(\mathcal{F}_t)$ *be the filtration generated by* W. *Show that the following processes are* $(\mathcal{F}_t)$*-martingales:*

(a) (W_t).

(b) $(W_t^2 - t)$.

(c) $(W_t^4 - 6tW_t^2 + 3t^2)$.

[*Hint: Note that* $W_t - W_s$ *is independent of* $\mathcal{F}_s$ $(0 \leq s \leq t)$.]

Exercise 6.20. *Let* (N_t) *be a Poisson process with parameter* $\lambda > 0$, *and put* $M_t = N_t - \lambda t$ $(t \geq 0)$. *Let* $(\mathcal{F}_t)$ *be the filtration generated by* (N_t).

(a) *Show that* (M_t) *is an* $(\mathcal{F}_t)$*-martingale.*
[*Hint:* $N_t - N_s$ *is independent of* $\mathcal{F}_s$ $(0 \leq s < t)$.]

(b) *Show that* $(M_t^2 - \lambda t)$ *is an* $(\mathcal{F}_t)$*-martingale.*
$[$*Hint: Write* $M_t^2 - M_s^2 = (M_t - M_s)^2 + 2M_s(M_t - M_s)$ $(0 \le s < t)$.$]$

Exercise 6.21. *Let* (N_t) *be a Poisson process with parameter* $\lambda > 0$*, and let* $c > 0$ *be any constant.*

(a) *Determine the constant* $\mu(c)$ *such that the process* $(\exp(cN_t + \mu(c)t))$ $(t \ge 0)$ *is a martingale with respect to the filtration* $(\mathcal{F}_t)$ *generated by* (N_t).
$[$*Hint: Use Theorem* 6.1.30 *and Exercise* 6.11.$]$

(b) *Verify directly that the process obtained in* (a) *is an* $(\mathcal{F}_t)$*-martingale..*
$[$*Hint: Use that* $\mathbb{E}[\exp(c(N_t - N_s))|\mathcal{F}_s] = \mathbb{E}[\exp(c(N_t - N_s))]$ *a.s.* $(0 \le s < t)$ *since* $N_t - N_s$ *is independent of* $\mathcal{F}_s$.$]$

Exercise 6.22. *Let* ξ *have a binomial distribution with parameters* $n \ge 1$ *and* $0 \le p \le 1$*, i.e.,*

$$P(\xi = k) \ = \ \binom{n}{k} p^k (1-p)^{n-k}, \ k = 0, 1, \ldots, n.$$

(a) *Use Azuma's inequality* (*Theorem* 6.3.1) *to prove the following inequality which is due to H. Chernoff* (*Ann. Math. Statist.* **23** (1952), $493-507$)*:*

$$P(|\xi - np| \ge t) \le 2\exp(-2t^2/n), \ t \ge 0, \ n \ge 1. \qquad (*)$$

$[$*Hint:* ξ *has the same distribution as a sum of suitable* $0-1$ *random variables* $\xi_1, \ldots, \xi_n$.$]$

(b) *Verify* $(*)$ *directly for* $n = 1$.

Exercise 6.23. *Prove* (6.147).
$[$*Hint: First note that* $|g(z)| =: \exp(I(z))$*, where*

$$I(z) \ := \ \int_0^z \frac{\cos x - 1}{x} \left(\log\left(\frac{z}{x}\right)\right)^r dx, \ z \ge 0, \ r > 0.$$

Then (6.147) *says that*

$$I(z) \le \frac{1}{2(r+1)} \left(1 - (\log(2z/(3\pi)))^{r+1}\right), \ z \ge 4\pi, \ r > 0. \qquad (*)$$

In order to prove $(*)$ *note that the cosine is* ≤ 0 *on the intervals* $J(k) := [(2k-1)\pi - \pi/2, (2k-1)\pi + \pi/2]$*, and that*

$$J(k) \subset [0, z] \quad \textit{iff} \quad 1 \le k \le k(z) \ := \ \lfloor z/(2\pi) + 1/4 \rfloor. \qquad (**)$$

Hence

$$I(z) \ \le \ -\sum_{k=1}^{k(z)-1} \int_{J(k)} \frac{1}{x} \left(\log\left(\frac{z}{x}\right)\right)^r dx.$$

Using $(**)$ *and comparing with a certain Riemann integral finally yields* $(*)$.$]$

Exercise 6.24. *A process* $Z_t = Z_0 \exp(X_t),\ t \geq 0\ (Z_0 > 0)$ *is observed at time points* $t = 0, 1, 2, \ldots, T$*, where* (X_t) *is a Lévy process of jump-diffusion type as in Example* 6.5.2*. Let* $H_0(2)$ *denote the null hypothesis which says that there exist* $\alpha \in \mathbb{R}, c \geq 2, \lambda \geq 0$ *and a distribution* Q_1 *on* $\mathbb{R}$ *satisfying* $Q_1(\{0\}) = 0$ *such that* (X_t) *is associated with* α, c, λ*, and* Q_1*.* (*Note that* $H_0(2)$) *has a meaning different from that at the beginning of* §6.5!) *Let* $H_0(2)$ *be rejected if* $|\widetilde{L}_T/T - p_{10}(1)| \geq 0.1$ (*see* (6.100) *and* (6.150))*. Let the level of significance be* 0.1*.* (*Note that the rejection of* $H_0(2)$ *entails the rejection of the null hypothesis that* (Z_t) *is a Black-Scholes process having volatility* ≥ 2 (*see* (6.27)).) *How large has* T *to be?* (*Answer:* $T \geq 1715$.)

Exercise 6.25. *A process* $Z_t = Z_0 \exp(X_t), t \geq 0\ (Z_0 > 0)$ *is observed at the time points* $t = 0, 1, 2, \ldots, T$*, where* $(X_t) = \alpha t + T_t, t \geq 0$*. Here,* $\alpha \in \mathbb{R}$; (T_t) *is a compound Poisson* (*or* CP-)*process associated with* $\lambda > 0$ *and* $Q_1 = N(\mu, \sigma^2)$ (*see Example* 6.1.21)*. Suppose that the null hypothesis* $H_0(\lambda^*, \sigma^*)$ $(\lambda^* > 0, \sigma^* > 0)$ *is to be tested, which says that there exist* $\alpha \in \mathbb{R}, \mu \in \mathbb{R}, \lambda \geq \lambda^*$*, and* $\sigma \geq \sigma^*$ *such that* $X_t = \alpha t + T_t$ $(t \geq 0)$*, and* (T_t) *is a* CP*-process associated with* λ *and* Q_1*. Verify that the test outlined in Exercise* 6.24*, which rejects* $H_0(\lambda^*, \sigma^*)$ *if* $|\widetilde{L}_T/T - p_{10}(1)| \geq 0.1$*, is not applicable no matter how the level of significance* $0 < p_0 < 1$ *is chosen.*
[*Hint: Show that there does not exist any* (*finite*) *constant* $\sum^*$ *satisfying* (6.153) (g *being the characteristic function of* X_1*,* (X_t)*being an arbitrary Lévy process satisfying* $H_0(\lambda^*, \sigma^*)$)*. Use Exercise* 6.14(b).]

Exercise 6.26. *Suppose we observe a process* $Z_t = Z_0 \exp(\mu t + cX_t), t \geq 0\ (Z_0 > 0)$ *at time points* $t = 0, 1, \ldots, T$*. Let* (X_t) *be a gamma process with parameters* α *and* Δ*, and consider* (*as in Example* 6.5.5) *the null hypothesis* $H_0(c^*, \alpha^*, \Delta^*)$ *where* $B = 10, c* = \alpha^* = 1, \Delta^* = 2, p_0 = v = 0.1, m = 1, d_1 = 1$*, and* $\lambda(10) = (2\pi/\log 10)^2$ (*recall that* $\log$ *is the natural logarithm*)*.*

(a) Show that in this special case we can choose $\sum^* = (\log 10)^2/24$*.*

(b) How large has the time horizon T *to be?* (*Answer:* $T \geq 2129$ (*instead of* $T \geq 2582$ *as in Example* 6.5.5!).)

Exercise 6.27. *Prove the following elementary result* (*Lemma* 6.6.7)*: Let* $a_1, a_2, \cdots$ *be real numbers such that* $0 \leq a_n < 1$ $(n \geq 1)$ *and* $\sum_{n=1}^\infty a_n < \infty$*. Then*

$$\sum_{n=1}^{\infty} a_n^t \to 0\ (t \to \infty).$$

Exercise 6.28. *Prove the claim in Example* 6.1.28*.*

Exercise 6.29. *Prove the iteration property of conditional expectations* (*see Example* 6.1.29).

Exercise 6.30. *Prove Lemma* 6.2.1*.*

PART 3
Applications I: Accounting and Vote Fraud

Chapter Seven

Benford's Law as a Bridge between Statistics and Accounting

Richard J. Cleary and Jay C. Thibodeau[1]

An auditor decides to run a Benford's law on a data set that consists of 1000 legitimate expense records from a business, plus a number of fraudulent transactions that an employee is making to a front for a business set up in a relative's name. Because the employees of the business have to obtain special approval for expenditures over $10,000, the fraudulent transactions are all for amounts between $9,000 and $9,999. For the 1000 legitimate expenditures, we have this data:

First Digit	Observed
1	314
2	178
3	111
4	92
5	88
6	59
7	56
8	56
9	46

Exercise 7.1. *Using the Benford law test at*

http://web.williams.edu/Mathematics/sjmiller/public_html/benford/chapter01/MillerNigrini_ExcelBenfordTester_Ver401.xlsx

(or any other suitable software), verify that the data conforms reasonably well to Benford's Law.

Exercise 7.2. *Use trial and error (or some more clever approach) to determine how many fraudulent transactions with first digit nine would need to be added to the 1000 legitimate observations above in order for the hypothesis that the data follows Benford's Law to be rejected at a five percent significance level. Does this seem plausible?*

Exercise 7.3. *What is the role of sample size in the sensitivity of Benford's law? Suppose there are 10,000 legitimate observations instead of 1000, but the ratios for legitimate observations remains the same, i.e., the number of observations for teach digit is multiplied by 10. Try the problem again. What changes?*

[1]Mathematics & Science Division, Babson College, and Department of Accountancy, Bentley University.

Exercise 7.4. *In which of the following situations is an auditor most likely to use Benford's Law?*

- *An analysis of a fast food franchise's inventory of hamburgers.*
- *An audit of a Fortune 500 company's monthly total revenue over the fiscal year.*
- *An analysis of a multi-billion dollar technology company's significant assets.*

Exercise 7.5. *Give an additional example of a way that including Benford's Law in an introductory-level statistics class will meet the four goals of the GAISE report of 2005.*

Exercise 7.6. *Determine whether the following situations are Type I errors, Type II errors, or neither.*

- *An auditor uses Benford's Law to analyze the values of canceled checks by a business in the past fiscal year. The auditor finds that there are significant spikes in the data set, with 23 and 37 appearing as the first two digits more often than expected. After further investigation, it was found that there were valid non-fraudulent explanations for the variations in the first digits.*
- *An auditor finds that a company's reported revenue does not follow Benford's Law. Further investigation is taken, and it is found that a manager has been rounding up her weekly sales to the nearest thousand to earn an incentive based on a weekly sales benchmark. The manager claims that the inflated sales were an accounting error.*
- *An owner of a business has falsely claimed to give his employees bonuses on each paycheck based on their monthly sales in order to lower his income taxes. An auditor examines the data, but is unable to confidently claim that the data does not follow Benford's Law. Rather than waste funds on a costly investigation, the auditor chooses not to investigate the owner.*

Exercise 7.7. *What are the negative effects of a Type I error in an audit? A Type II error? In what situations might one be more dangerous than the other?*

Exercise 7.8. *What are some of the reasons listed in the chapter that might explain why a data set should not be expected to follow Benford's Law?*

Exercise 7.9. *Give an example of a reason other than fraud that explains why a data set that is expected to conform to Benford's Law does not.*

Chapter Eight

Detecting Fraud and Errors Using Benford's Law

Mark Nigrini[1]

Exercise 8.1. *Do the following data sets meet the requirements described by Nigrini in order to be expected to follow Benford's Law? Explain why or why not.*

- *The 4-digit PIN numbers chosen by clients of a local bank.*
- *The annual salaries of graduates from a public university.*
- *Numeric student ID numbers assigned by a school.*
- *The distances in miles between Washington, DC and the 500 most populated cities in the United States (excluding Washington, DC).*
- *Results to a survey of 1,000 students asked to provide a number in between 1 and 1,000,000.*
- *The number of tickets bought for all events held in a particular stadium over the past five years.*

Exercise 8.2. *Take a company which has been at the heart of a scandal (for example, Enron) and investigate some of its publicly available data.*

Exercise 8.3. *An audit of a small company reveals a large number of transactions starting with a 5. Come up with some explanations other than fraud.* Hint: there are two cases: it is the same amount to the same source each time, and it isn't.

[1]Department of Accounting, West Virginia University, Morgantown, West Virginia 26506.

Chapter Nine

Can Vote Counts' Digits and Benford's Law Diagnose Elections?

Walter R. Mebane, Jr.[1]

Exercise 9.1. *If X satisfies Benford's law, then the mean of its second digit is 4.187. What is the mean of the k^{th} digit?*

Exercise 9.2. *If X satisfies Benford's law, multiply by an appropriate power of 10 so that it has k integer digits. What is the probability the last digit is d? What is the probability the last two digits are equal? What is the probability the last two digits differ by 1?*

Exercise 9.3. *Find some recent voting data (say city or precinct totals) and investigate the distribution of the first and second digits.*

[1]Department of Political Science and Department of Statistics, University of Michigan, Ann Arbor, MI. The author thanks Jake Gatof, Joe Klaver, William Macmillan and Matthew Weiss for their assistance.

Chapter Ten

Complementing Benford's Law for small N: a local bootstrap

Boudewijn F. Roukema[1]

Exercise 10.1. *Do you agree with the assessment that Nigrini's conditions for applying Benford's Law are mostly satisfied? Why or why not?*

Exercise 10.2. *Why does having a large $\sigma(\log_{10} x_i)$ and a large $\sigma(\log_{10} w_{i,j})$ ensure that $v_{i,j}$ first-digit distribution approaches Benford's Law?*

Exercise 10.3. *What does it mean for bootstrap methods to be considered "conservative?" Identify some of the ways in which bootstrap methods are conservative.*

Exercise 10.4. *There are many conservative statistics. Look up the Bonferroni adjustment for multiple comparisons, as well as alternatives to that.*

Exercise 10.5. *How would a local bootstrap realization change if the value of Δ were changed?*

Exercise 10.6. *Confirm that if $c_{bK7} > 99.924\%$, then $c_{eK7} > 99.99960\%$.*

[1]Toruń Centre for Astronomy, Faculty of Physics, Astronomy and Informatics, Nicolaus Copernicus University, ul. Gagarina 11, 87-100 Toruń, Poland.

PART 4
Applications II: Economics

Chapter Eleven

Measuring the Quality of European Statistics

Bernhard Rauch, Max Göttsche, Gernot Brähler, Stefan Engel[1]

Exercise 11.1. *In which of the following two scenarios would χ^2 be larger?*

- *The first-digit frequencies are mostly identical to the expected Benford distribution, but the digit 1 appears 31.1% of the time and the digit 2 appears 16.6% of the time (compared with the expected values of approximately 30.1% and 17.6%, respectively)*
- *The first-digit frequencies are mostly identical to the expected Benford distribution, but the digit 8 appears 6.12% of the time and the digit 2 appears 3.58% of the time (compared with the expected values of approximately 5.12% and 4.58%, respectively)*

Exercise 11.2. *What is μ_b, the value of the mean of the Benford distribution of first digits base b?*

Exercise 11.3. *What is the value of a^* if μ_e = 3.5?*

Exercise 11.4. *Using Figure 11.1, confirm the values of χ^2, χ^2/n, and d^* for the distribution of first digits for Greece social statistics in the year 2004.*

Exercise 11.5. *Using Figure 11.1 and the formula for distance measure a^* used by Judge and Schechter, calculate the value of the mean of the dataset (μ_e) in the year 2004. Confirm this value by using the formula $\mu_e = \frac{\Sigma_{i=1}^{9} n\text{Prob}(D_1=i)}{n}$.*

The final problem uses data on two fictitious countries, which is available online

http://web.williams.edu/Mathematics/sjmiller/public_html/benford/chapter11/

(some of the additional readings on that webpage may be useful as well).

Exercise 11.6. *Calculate the values χ^2, χ^2/n, d^* and a^* and compare the results for both countries. Which one of these two countries should be examined closer? Are the outcomes consistent?*

[1]Rauch: University of Regensburg, Department of Economics, Universitätsstraße 31, 93053 Regensburg, Germany; Göttsche and Engel: Catholic University of Eichstätt-Ingolstadt, Department of Auditing and Controlling, Auf der Schanz 49, 85049, Ingolstadt, Germany; Brähler: Ilmenau University of Technology, Department of Taxation Theory and Auditing, Helmholtzplatz 3, 98693 Ilmenau, Germany.

Chapter Twelve

Benford's Law and Fraud in Economic Research

Karl-Heinz Tödter[1]

Exercise 12.1. *Use (12.1) to find $f(6)$ and $F(6)$ for Benford's Law.*

Exercise 12.2. *If X is a Benford variable defined on $[1, 10)$, then what is the probability that the second digit is 5 given that the first digit is also 5?*

Exercise 12.3. *Use (12.4) to confirm that when using Benford's Law for Rounded Figures, $Prob(D_1 = 8) = 0.054$.*

Exercise 12.4. *If X is a Benford variable defined on $[1, 10)$, given that the first digit is 8, what is the probability that the second digit is 0 when rounding to two significant digits? What is the probability that the second digit is 2?*

Exercise 12.5. *Using Benford's Law for Rounded Figures as the frequencies of first digits for a data set of 300 observed values, calculate Q_1, Q_2, M_1, and M_2 using (12.6) and (12.7).*

Exercise 12.6. *Should the Q_1 test or the M_1 test be used for attempting to detect variations in Benford's Law?*

- *What if the data set in question has a mean of 3.44?*
- *Which test should be used for detecting variations in the Generalized Benford's Law?*

Exercise 12.7. *The Federal Tax Office (FTO) knows that $\Omega = 10\%$ of tax declarations of small and medium enterprises are falsified. The FTO checks the first digits using Benford's Law. Random samples of tax declarations are drawn and the null hypothesis ($\mathrm{H_o}$) "Conformity to Benford's Law" is tested at the $\alpha = 5\%$ level of significance.*

- *Using (12.9), what rejection rate of $\mathrm{H_o}(\theta)$ would you expect if the probability of a type II error β lies in the interval [0.05, 0.75]?*
- *The FTO obtained the rejection rate $\theta = 0.12$. Use (12.9) to calculate the probability β of a type II error.*
- *The FTO arranges for an audit at the taxable enterprise if the Benford test rejects $\mathrm{H_o}$ for a certain tax declaration at the $\alpha = 5\%$ level. What is the probability that such an audit will be provoked erroneously? And what is the probability to forbear an audit erroneously?*

[1]Research Centre, Deutsche Bundesbank

Exercise 12.8. *A sample of scientific articles is taken, and 17% are found to have regression coefficients with a doubtful distribution of first digits. Use (12.10) to calculate* $\hat{\Omega}$.

Chapter Thirteen

Testing for Strategic Manipulation of Economic and Financial Data

Charles C. Moul and John V. C. Nye[1]

Exercise 13.1. *What are some of the potential reasons given in Section 13.1 for why data sets that are expected to follow Benford's Law fail to do so?*

Exercise 13.2. *Did Benford's Law prove financial misreporting during the financial crisis? Justify your assertion.*

Exercise 13.3. *What are some of the potential motives that banks have for manipulating VAR data?*

[1] Economics Department, Farmer School of Business, Miami University, Ohio and Mercatus Center and Economics Department, George Mason University and National Research University – Higher School of Economics, Moscow, respectively; We thank Marc Taub for excellent research assistance.

PART 5
Applications III: Psychology and the Sciences

Chapter Fourteen

Psychology and Benford's Law

Bruce D. Burns and Jonathan Krygier[1]

Exercise 14.1. *Using (11.1) in Section 11.3, find* χ^2 *for the elaborated and unelaborated data from Scott, Barnard and May's study found in Table 14.1.*

Exercise 14.2. *What distribution of leading digits would you expect if people were asked to randomly give an integer from* 1 *to* N*? How does your answer depend on* N*? Try an experiment with some of your friends and family.*

[1]School of Psychology, The University of Sydney, NSW 2006, Australia. The authors would like to thank to Hal Willaby for comments on an earlier draft. The authors were supported by a grant from the University of Sydney, and it is a pleasure to thank them for their generosity.

Chapter Fifteen

Managing Risk in Numbers Games: Benford's Law and the Small Number Phenomenon

Mabel C. Chou, Qingxia Kong, Chung-Piaw Teo and Huan Zheng[1]

Exercise 15.1. *What are the risks associated with a high liability limit in a fixed-odds lottery game? What if the limit is too small?*

Exercise 15.2. *From the data obtained in Table 15.1, determine the probability that a given number on a ticket for the UK powerball game is a single digit.*

Exercise 15.3. *Figure 15.1 shows the proportion of tickets in a Pennsylvania Pick-3 game with a given first digit. Explain why there are several outliers larger than the mean proportion and no outliers smaller than the mean proportion.*

Exercise 15.4. *What is the probability that a Type I player chooses the number 345 in a Pick-3 game?*

Exercise 15.5. *Let Alice be a Type II player in a Pick-3 game that bets on a number with three significant digits 80% of the time, a number with two significant digits 15% of the time, and a number with one significant digit 5% of the time. What is the probability that Alice bets on the number 345? The number 45? The number 5?*

Exercise 15.6. *In the Pennsylvania Pick-3 game, the least square model indicates that 60.42% of the players are Type I players and 39.58% of the players are Type II players. Based on this model, use (15.4) to calculate the expected proportion of the betting volume on a three-digit number with first significant digit 4.*

Exercise 15.7. *Let Bob be a Type II player in a Pick-3 game that bets on a number with three significant digits 80% of the time, but also has a tendency to exhibit switching behavior; that is, he will switch later digits with probability 0.9105, and switch the digit to 0 with probability 0.1054. What is the probability that Bob bets on the number 345?*

Exercise 15.8. *Use (15.5) to calculate the probability that Bob chooses a three-digit number in between 520 and 529 inclusive.*

Exercise 15.9. *Calculate the variance using the equation in 15.4.1 under the scenario that all players randomly select a three-digit number.*

[1] Chou and Teo: National University of Singapore; Kong: Universidad Adolfo Ibañez; Zheng: Shanghai Jiao Tong University.

Chapter Sixteen

Benford's Law in the Natural Sciences

David Hoyle[1]

Exercise 16.1. *Demonstrate that (16.3) holds for* $\alpha = 2$.

Exercise 16.2. *Rewrite the lognormal distribution density function (16.5) as the lognormal density function (16.6).*

Exercise 16.3. *Show that as* σ *grows larger, the lognormal density function approaches the power law* $p(x) = C_\sigma x^{-1}$, *where* C_σ *is a constant depending on* σ.

Exercise 16.4. *Provide examples not mentioned in the chapter of scientific data sets that are not effectively scale-invariant.*

Exercise 16.5. *Explain the intuition behind why the following distributions are approximately Benford:*

- *The Boltzman-Gibbs distribution (16.8).*
- *The Fermi-Dirac distribution (16.9).*
- *The Bose-Einstein distribution (16.10).*

Exercise 16.6. *Obtain a physics textbook (or a CRC handbook, or...) and find a list of physical constants. Perform a chi-squared test to determine if the list of constants follow Benford's Law as expected.*

Exercise 16.7. *Sandon found agreement between Benford's Law and population and surface area data for the countries of the world. Find a source that provides the population density of each country. Then determine if population density follows Benford's Law. This can be done using a chi-squared test. In general, should the ratio of two of two Benford random variables be Benford?*

[1] Thorpe Informatics Ltd., Adamson House, Towers Business Park, Wilmslow Rd., Manchester, M20 2YY, UK.

Chapter Seventeen

Generalizing Benford's Law: A Re-examination of Falsified Clinical Data

Joanne Lee, Wendy K. Tam Cho, and George Judge[1]

Exercise 17.1. *Use (17.1) to calculate the average frequency of first digits in Stigler's distribution of first significant digits. Check to see that the distribution matches the values displayed in Table 17.1.*

Exercise 17.2. *Verify (17.3), (17.5), and (17.6). Then verify that the sum of the three subsets matches (17.7).*

Exercise 17.3. *Calculate the mean of the Stiegler FSD distribution and Benford FSD distribution to confirm that they are equivalent to 3.55 and 3.44, respectively.*

Exercise 17.4. *For the Estimated Maximum Entropy FSD distribution for data with a FSD mean of 3.44 shown in Table 17.3, find $H(p)$ and ensure that the criterion from (17.13) and (17.14) are reached.*

- *If the Estimated Maximum Entropy FSD distribution is accurate, then the listed probabilities will maximize $H(p)$. First, determine if replacing $\hat{p_1}$ with 0.231 and $\hat{p_2}$ with 0.2 still allows (17.13) and (17.14) to hold. Now find $H(p)$. Is $H(p)$ larger or smaller than before?*

Exercise 17.5. *If the FSD mean is 5, what will be the estimated maximum entropy FSD distribution? What is $Var(d)$ according to (17.18)?*

Exercise 17.6. *Examining the Poehlman data in Table 17.4, calculate the difference for each digit FSD distribution given by Benford's Law.*

Exercise 17.7. *The estimated empirical likelihood distributions given a FSD mean will maximize $\sum_{i=1}^{9} p_i$. To test this, ensure that the product of the $p_i's$ from Table 17.5 are greater than the empirical data found in Table 17.4.*

Exercise 17.8. *A researcher is trying to decide if a dataset follows Benford's law or Stigler's law. What values of the mean of the leading digit suggest Benford over Stigler? What values suggest Stigler over Benford?*

[1]Lee: Researcher, Mathematica Policy Research; Cho: Departments of Political Science and Statistics, and Senior Research Scientist, National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign; Judge: Department of Agricultural and Resource Economics, University of California at Berkeley.

PART 6

Applications IV: Images

Chapter Eighteen

Partial Volume Modeling of Medical Imaging Systems using the Benford Distribution

John Chiverton and Kevin Wells[1]

Exercise 18.1. *What is the PV effect? What implications does the PV effect have for medical imaging?*

Exercise 18.2. *Prove Corollary 18.3.4.*

Exercise 18.3. *What advantages are there to describing the PV effect using matrices as in (18.11)?*

Exercise 18.4. *What are the differences between a Rician noise model described by (18.12) and a Gaussian noise model described in (18.13)?*

Exercise 18.5. *Use (18.22) to calculate* $p(\alpha)$ *for* $\alpha = 0.50$*, where* α *has two digits of precision.*

Exercise 18.6. *How is the contrast to noise ratio (CNR) affected if both the distance between the signal levels of two components and the standard deviation of each class is doubled.*

[1] J. Chiverton is with the School of Engineering, University of Portsmouth, UK and K. Wells is with the Centre for Vision, Speech and Signal Processing, University of Surrey, UK. *john.chiverton@port.ac.uk, k.wells@surrey.ac.uk*.

Chapter Nineteen

Application of Benford's Law to Images

Fernando Pérez-González, Tu-Thach Quach, Chaouki T. Abdallah, Gregory L. Heileman and Steven J. Miller[1]

Exercise 19.1. *In (19.9) one of the factors is* $\Gamma\left(\frac{-j2\pi n+\log 10}{c\log 10}\right)$*, where* $j=\sqrt{-1}$*. Estimate how rapidly this tends to zero as* $|n|\to\infty$ *as a function of* c *(if you wish, choose some values of* c *to get a feel of the behavior).*

Exercise 19.2. *In (19.19) we find that* $|a_n(c,\sigma)|\le|a_n(c^+)|$ *for all* n*; investigate how close these can be for various choices of* c *and* σ*.*

Exercise 19.3. *In Example 19.5.1 we found four zero-mean Gaussians with shaping parameter* $c=1$ *with four different standard deviations and* $a_1=0$*. Can you find six zero-mean Gaussians with shaping parameter* $c=1$ *and six different standard deviations with* $a_1=0$*? What about eight? More generally, can you find* $2m$ *such Gaussians for* m *a positive integer?*

[1]Perez-Gonzalez: Department of Signal Theory and Communications, University of Vigo, EE Telecomunicacion, Campus Universitario, 36310 Vigo, Spain; Quach: Sandia National Laboratories, Albuquerque, NM; Miller: Department of Mathematics and Statistics, Williamstown, MA; Abdallah and Heileman: Electrical & Computer Engineering Department, University of New Mexico, Albuquerque, NM.