AN ELEMENTARY PROOF OF STIRLING'S FORMULA

BY

P. DIACONIS AND D. FREEDMAN

TECHNICAL REPORT NO. 37
REVISED
DECEMBER 1984

RESEARCH PARTIALLY SUPPORTED BY
NATIONAL SCIENCE FOUNDATION
GRANTS MCS80-24649 AND MCS83-01812

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA
December, 1984
To appear in American Math. Monthly

AN ELEMENTARY PROOF OF STIRLING'S FORMULA

by

P. Diaconis\textsuperscript{1}
Statistics Department
Stanford University

and

D. Freedman\textsuperscript{2}
Statistics Department
University of California, Berkeley

Abstract. The object of this note is to give an elementary proof of Stirling's formula, using a direct from of Laplace's method.

\textsuperscript{1}Research partially supported by National Science Foundation Grant MCS80-24649.

\textsuperscript{2}Research partially supported by National Science Foundation Grant MCS83-01812.

AMS/MOS classification numbers. 33A15, 60F05.

Keywords and phrases. Stirling's formula, factorials, normal approximation.

Running head. Stirling's formula.
1. Introduction. Stirling's formula is

\[ \Gamma(\alpha) \approx \left(\frac{\alpha-1}{e}\right)^{\alpha-1} \sqrt{2\pi(\alpha-1)} \]

as \( \alpha \to \infty \), in the sense that the ratio of the two sides tends to 1. By definition,

\[ \Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x} \, dx \quad \text{for} \quad \alpha > 0. \]

Thus, for positive integers \( n \), \( \Gamma(n) = (n-1)! \).

A formal argument for (1) is presented in the next section, but here is the idea. Write \( \exp(x) = e^x \) and \( \beta = \alpha - 1 \). Then

\[ \Gamma(\alpha) = \int_0^\infty \exp(\psi_\beta(x)) \, dx \]

where

\[ \psi_\beta(x) = \beta \log x - x. \]

Now \( \psi_\beta \) has its maximum at \( x = \beta \), and

\[ \psi_\beta(\beta+y) \doteqdot \beta \log \beta - \beta - \frac{1}{2\beta} y^2. \]

Here, \( \doteqdot \) means "nearly equal," and is used informally. So

\[ \Gamma(\alpha) \doteqdot \exp(\beta \log \beta - \beta) \star \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2\beta} y^2\right) \, dy \]

\[ = \left(\frac{\beta}{e}\right)^\beta \sqrt{2\pi\beta} \]

because

\[ \int_{-\infty}^{\infty} \exp(-\frac{1}{2}z^2) \, dz = \sqrt{2\pi}. \]
2. The argument. Recall \( \psi_\beta \) from (4). The rigorous version of (5) is the following identity:

\[
(8) \quad \psi_\beta(\beta+y) = \beta \log \beta - \beta - \beta g(y/\beta)
\]

where

\[
(9) \quad g(v) = v - \log(1+v).
\]

Substitute (8) into (3) and change variables:

\[
(10) \quad \Gamma(\alpha) = \left( \frac{\beta}{e} \right)^\beta \int_{-\beta}^{\beta} \exp\{-\beta g(y/\beta)\} \, dy.
\]

Change variables again, putting \( y = \sqrt{\beta} z \),

\[
(11) \quad \Gamma(\alpha) = \left( \frac{\beta}{e} \right)^{\frac{\beta}{2\sqrt{\beta}}} \sqrt{2\pi \beta} \Gamma_1(\beta)
\]

where

\[
(12) \quad \Gamma_1(\beta) = \frac{1}{\sqrt{2\pi}} \int_{-\sqrt{\beta}}^{\sqrt{\beta}} \exp\{-\beta g(z/\sqrt{\beta})\} \, dz.
\]

Our proof of Stirling's formula is reduced to the following.

**Lemma.** \( \lim_{\beta \to \infty} \Gamma_1(\beta) = 1 \)

**Proof.** Fix \( L \) large but finite. Then \( \Gamma_1(\beta) = \Gamma_L(\beta) + \tau_L(\beta) \)

where

\[
(13) \quad \Gamma_L(\beta) = \frac{1}{\sqrt{2\pi}} \int_{-L}^{L} \exp\{-\beta g(z/\sqrt{\beta})\} \, dz
\]

\[
\to \frac{1}{\sqrt{2\pi}} \int_{-L}^{L} \exp\{-\frac{1}{2} z^2\} \, dz
\]
because

\[ \beta g(z/\sqrt{\beta}) \to \frac{1}{2} z^2 \quad \text{as} \quad \beta \to \infty, \text{uniformly for} \quad z \in [-L, L]. \]

Relationship (14) holds because for \( v \) small enough,

\[ (1-\varepsilon)\frac{1}{2} v^2 < g(v) < (1+\varepsilon)\frac{1}{2} v^2 \]

It remains to estimate \( \tau_L(\beta) \). We take the upper tail first. Abbreviate \( h(z) = \beta g(z/\sqrt{\beta}) \); the dependence of \( h \) on \( \beta \) is implicit. Then

\[ \int_L^\infty \exp\{-h(z)\} \, dz \leq \frac{1}{h'(L)} \int_L^\infty h'(z) \exp\{-h(z)\} \, dz \]

\[ = \frac{1}{h'(L)} \exp\{-h(L)\} \]

\[ + \frac{1}{L} \exp\{-\frac{1}{2} L^2\} \quad \text{as} \quad \alpha \to \infty; \]

the first inequality holds because \( h'(z) = \sqrt{\beta} z/(\sqrt{\beta} + z) \) is increasing in \( z > 0 \). The lower tail is similar, so for \( \beta \) large,

\[ \tau_L(\beta) \leq \frac{2}{\sqrt{2\pi}} \frac{1}{L} \exp\{-\frac{1}{2} L^2\} + \varepsilon \]

The present approximation comes out in powers of \( \beta = \alpha - 1 \); if powers of \( \alpha \) are preferred, a preliminary change of variables can be made \( (x = e^u) \), so

\[ \Gamma(\alpha) = \int_{-\infty}^{\infty} u^\alpha \exp(-e^u) \, du. \]

A refinement of the argument leads to the usual asymptotic development,

\[ \Gamma_1(\beta) = 1 + \frac{1}{12\beta} - \ldots \]
3. History. The first proofs of Stirling's formula were given by de Moivre (1730) and Stirling (1730). Both used what is now called the Euler-MacLaurin formula to approximate $\log 2 + \log 3 + \ldots + \log n$. De Moivre proved the result on the way to the normal approximation for the binomial distribution. His first derivation did not explicitly determine the constant $\sqrt{2\pi}$. In a 1731 addendum, he acknowledged that Stirling was able to determine the constant, using Wallis' formula. To statisticians, the most familiar version of this argument is Robbins (1955); also see Feller (1968). The approach yields upper and lower bounds for $n!$ but does not extend to $\Gamma(a)$.

In essence, we are using a direct form of Laplace's (1774) method to estimate the integral in (12), with quite explicit bounds. For a more general treatment, see de Bruijn (1981, sec. 4.5). For one very similar to ours, see Woodroffe (1975, p. 127). We know of three other approaches to proving Stirling's formula. Modern analysts extend $\Gamma$ into the complex plane, and have a proof of (1) using the saddlepoint method: see de Bruijn (1981, sec. 6.9). Artin (1964) presents a fascinating discussion of the $\Gamma$-function and its properties, as well as a proof of Stirling's formula based on the following theorem: $\Gamma(a)$ is the only log convex function on $(0, \infty)$ satisfying $\Gamma(a+1) = a\Gamma(a)$, with $\Gamma(1) = 1$. A third approach using the residue calculus is due to Lindelöf; for a modern exposition, see Ahlfors (1979).

We stumbled on our proof while working on finite forms of de Finetti's theorem for exponential families (Diaconis and Freedman, 1980, 1984). As an example, we were thinking of the gamma shape parameter $\alpha > 0$ in the family

$$\frac{1}{\Gamma(a)} x^{a-1} e^{-x} \quad \text{for} \quad 0 \leq x < \infty.$$
As $\alpha \to \infty$, the density (20) tends to normal, with mean $\beta$ and variance $\beta$. The argument for Stirling's formula was a by-product.
REFERENCES


