Artificial Intelligence
for Human-Technologies-
Economy Sustainable
Development
23–25 May 2024
Lublin, Poland

Management,
Knowledge and Learning
International Conference
Technology, Innovation
and Industrial Management
Sustainable Development

Why is There π in the Normal Distribution? Application of Reductionism Principles to Explain a Pure Mathematical Formula

Fabio Pisano
Pegaso International, Malta
pisano.fabio@hotmail.it

Lidia Nazzaro
Pegaso International, Italy
lidianazzaro@virgilio.it

Abstract

A proper usage of any formula requires a full understanding of its meaning. The presence of some famous mathematical constant (such as π or e) is not straightforward and it can be quite a lengthy process explaining all the steps of a mathematical demonstration if not all the theorems or the formula adopted are widely known. It is possible to answer the question without providing an exhaustive explanation of all the steps to reach such a formula. The answer for the title is the following: Application of trigonometric and calculus to a combinatoric function.

$$\mathbb{N}(\mu,\sigma) = \frac{e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}{\sqrt{2\boldsymbol{\pi}\sigma^2}}$$

Keywords: π, calculus, trigonometrics, combinatoric, reductionism

1. INTRODUCTION

There is a story about two friends, who were classmates in high school, talking about their jobs. One of them became a statistician and was working on population trends. He showed a reprint to his former classmate. The reprint started, as usual, with the Gaussian distribution, and the statistician explained to his former classmate the meaning of the symbols for the actual population, for the average population, and so on. His classmate was a bit incredulous and was not quite sure whether the statistician was pulling his leg. "How can you know that?" was his query. "And what is this symbol here?" "Oh," said the statistician, "this is pi." "What is that?" "The ratio of the circumference of the circle to its diameter." "Well, now you are pushing your joke too far," said the classmate, "surely the population has nothing to do with the circumference of the circle."[1]

Whenever any individual refers for any purpose to the Normal Distribution, one of the reply own opponent may state is the following: Why are you using “π” in the normal distribution?

From a pure mathematical point of view this question is meaningless, although consider the following sillogism:

- **π** is a constant present in the so-called normal distribution, which is a very important equation from a Statistical point of view;
- The applications of the so-called normal distribution may be very relevant from a practical point of view;
- Indeed, **π** may have very relevant consequences from a practical point of view;

Actually, such equation is available since XVIII century (from the time of Abraham de Moivre).

Such exercise is more relevant for those individuals not working in pure mathematical field, since such concept is not very straightforward.

Consider, for example, the incomes of the citizens of two cities, a variable which is consider to to fit the “normal distribution”. In both cases values seem to fit the so-called “Normal Distribution” and this information is willing to be used by the fiscal bureau, reducing my fiscal burden. In this very simple example, such information is not just a mathematical concept, but it is attached also to a monetary flow: hence it turns a relevant information.

If we want to adopt such fiscal policy also into our city, we need not only the mathematical result, but also understanding what it represents and why, for example, the same mathematical result can lead to two different consequences in two different samples.

A possible informal answer to such question is the following:

The presence of the mathematical constant π, as well as the mathematical constant “e” into the formula of the Normal Distribution is due to the fact such function is obtained through the application of trigonometrics and the calculus to a combinatoric function.

Although, trying to deepening the explanation, we had found that there are a series of mathematical theorems (Taylor Series, Theorem of Fundamental Algebra, Squeeze Theorem, Trigonometric identities) which are involved into such demonstration and we are not providing them their proofs: this is a limit of using the reductionism from a reductionist as a tool to provide an exhaustive

---

[1] https://www.maths.ed.ac.uk/~v1ranick/papers/wigner.pdf

explanation of such mathematical constant (since we do know why π appears in the Normal Distribution without knowing some steps of its demonstration).

2. COMBINATORICS, THE BINOMIAL DISTRIBUTION

$$\sum_{k=0}^{n} \binom{n}{k} * p^{(k)} * q^{(n-k)} = \frac{n!}{(n-k)! * k!} * p^{(k)} * q^{(n-k)}$$

Combinatorics is an ex-post type of reasoning:

The only thing we know, or the only thing we are interested in, are the actual results of a specific process, not the forces which lead to it.

Consider the very famous example of rolling a dice. There is a plethora of forces in action (smoothness of the surface, forces of my hand, elasticity of the dice), although the possible results are just 6 (the number of the face of the dice).

A similar condition occurs in many other instances: for example, a modern variation of the idea of extracting black and white balls from a jar (an example used frequently by Bernoulli in his *Ars Conjectandi*) is the analysis of the parking area for 10 places of an office.

At this stage, we can object that each driver is different one each other (different driving style, different car, different time of departing, different road with different traffic). How can we pretend that combinatorics may be applied in such case?

What is said is true. Although it is not the full story since we may not be interested in knowing the dynamics of all the determinants: If each single contribution (or a combination of them) may be outperformed by the rest of the others, we can have a "Robust Distribution", hence a distribution in which removing one element into our analysis does not affect the results.[2]

For example: if there is no flexibility at entrance (parking area opens at 8 a.m.), slower drives can simply depart earlier and still being able to find a parking space inside. A further objection may come from "variance": two employee use two different roads respectively, one of these road has a constant flux of traffic, the other can heavily change.

Nevertheless, it does not affect the genuinity of the reasoning: Is it possible knowing in which day the road has more traffic or not?

- If not, it implies that the combinatoric reasoning can be applied to the traffic road too;
- If yes, it may be internalized by the drivers,[3] otherwise it effectively affects the genuinity of such reasoning.

In this concern, there is room for a correct analysis, in such case, the binomial coefficient must be pondered by two probabilities "p" (the probability of finding a parking space inside the office) and "1-

[2] This is may due for two reasons:
1) because there are so many causes which are interacting one each others, the combination of which is not ordered;
2) because a "small variation" beyond our measurement has a relevant effect into the distribution which can be measured.
Actually, the latter is the definition of "deterministic chaos" (popularly called "butterfly effect") whilst the formed is a definition of "stochastic distribution", they are two different concepts.

[3] For example, if it is known that each Tuesday the road has more traffic, if the employee has such possibility he could depart earlier so that he will manage to arrive on time.

p = q” (the complementary case, hence the probability of not finding a parking space inside the office) which can be obtained through the frequentist reasoning: p = 10/15, q = (15-10)/15.

Each week I have the possibility of parking inside the office or outside 5 times, hence: “n” = 5. If I park outside I will pay 1 Euro per day. How much can I spend every week?

Thanks to a computer, we can easily compute all the possible combinations, in such case, we will have a dataframe like this one. There are 32 possible permutations, hence 2^5 Possible results.

CHART 1

Monday	Tuesday	Wednesday	Thursday	Friday	**EURO Spent**
0	0	0	0	0	**0**
1	0	0	0	0	**1**
0	0	1	0	0	**1**
0	1	0	0	0	**1**
0	0	0	1	0	**1**
0	0	0	0	1	**1**
1	1	0	0	0	**2**
0	0	0	1	1	**2**
0	0	1	0	1	**2**
0	0	1	1	0	**2**
1	0	1	0	0	**2**
1	0	0	1	0	**2**
1	0	0	0	1	**2**
0	1	0	0	1	**2**
0	1	0	1	0	**2**
0	1	1	0	0	**2**
1	1	0	0	1	**3**
0	0	1	1	1	**3**
0	1	0	1	1	**3**
0	1	1	0	1	**3**
0	1	1	1	0	**3**
1	1	1	0	0	**3**
1	0	0	1	1	**3**
1	0	1	0	1	**3**
1	0	1	1	0	**3**
1	1	0	1	0	**3**
1	1	0	1	1	**4**
1	1	1	0	1	**4**
1	1	1	1	0	**4**
1	0	1	1	1	**4**
0	1	1	1	1	**4**
1	1	1	1	1	**5**

This method, although is very time consuming, In fact, if we increase the numerosity to a month 2^{22} = 4,194,305.

If we try one year (roughly 300 days per year), 2^{300} (around 10^{90}) it will be too time consuming and too inefficient to compute with such algorithm.[4]

If we add the other employee of the office in the reasoning, for example 15, we can go further. Every day, the first employee can choose between 10 possible parking space, the second between 9, until the

[4] We can track it through the time of processing, as well as the size of the dataframe occupied.

tenth, who will find the latest parking space. This represents a product spanning from “n-k: number of cars which should be parked outside”, until “n: all the cars”

$$\frac{n!}{k!} = \frac{\cancel{1*2*3*\ldots*(k-1)*k}*(k+1)*\ldots n}{\cancel{1*2*3*\ldots*(k-1)*k}} = \prod_{i=k+1}^{n} i$$

In such case, each parking distribution inside the office is associated with several parking distribution outside the office. However, as long as we are studying what is happening just inside the parking of the office, we are willing to remove these cases; hence, we are going to remove all the “(n-k)!” from the product above, obtaining the binomial coefficient.

$$\frac{n!}{k!\,(n-k)!}$$

Hence, pondering the binomial coefficient with the frequencies “p” and “q”, we can easily get all these steps, and we can obtain the binomial distribution:

$$\frac{n!}{(n-k)!*k!}*p^{(k)}*q^{(n-k)} \to \sum_{k=0}^{n}\binom{n}{k}*p^{(k)}*q^{(n-k)} = \left(1+\frac{1}{r}\right)^{n}$$

At this stage, from a computational point of view, we are stuck due to the high cardinality of the numbers involved. The way in which this problem was tackled was through the concept of factorial “n!”. To overcome this problem, there are a couple of steps which are involved,

The binomial distribution shown above is a polynomial function,

$$f(x) = \sum_{i=1}^{n} a_i\, x^i$$

we can take advantage of it using the Viete Theorem. Moreover, it is necessary to use the Stirling Approximation formula:

$$n! = \sqrt{2\boldsymbol{\pi} n} * (\frac{n}{e})^n$$

Here, we clearly identify the presence of both mathematical constants “π” and “e”. However, for the purpose of this article, we are just moving forward with the question:

Why is there “π” in the Stirling formula?

The answer is:

Because the Stirling formula can be referred to as the Wallis Product:

$$\prod_{n=1}^{\infty}\frac{2n}{2n-1}*\frac{2n}{2n+1} = \frac{\pi}{2}$$

So again:

Why is there "π" in the Wallis formula?

The next chapter will provide such steps.

In the meantime, an important element is the "inference problem" with "inference process", a logical process that allows the extension of the properties of a determined set (the sample) to an indefinite one (the population).

Once this is given, we can justify the following transformation:

$$f(x) = \frac{n!}{n! * k!} * p^{(k)} * q^{(n-k)} = \frac{\sqrt{2\pi n} * (\frac{n}{e})^n}{\sqrt{2\pi(n-k)} * (\frac{n-k}{e})^{n-k} * \sqrt{2\pi(n-k)} * (\frac{k}{e})^k} * p^{(k)} * q^{(n-k)}$$

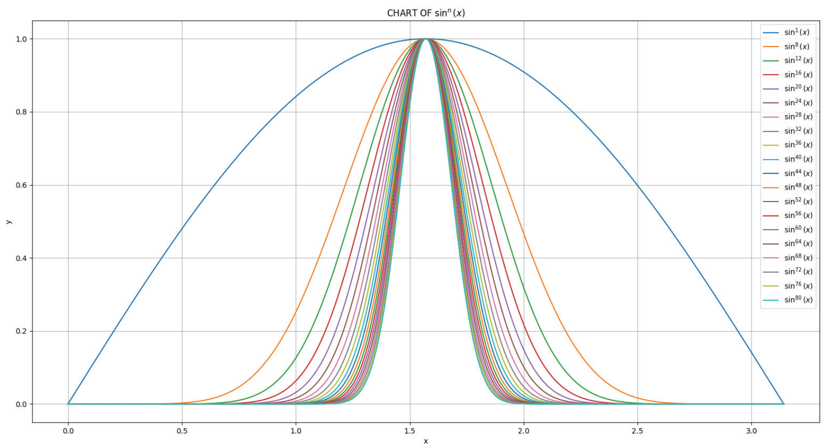
Hence, following Hald (chapter 14 and chapter 24) it is possible obtaining the Poisson an the Normal Distribution.

Even when the nature of the formula is changed, it is possible to reason based on a limited number of observations. Numbers, although, may increase indefinitely.

Indeed, it is possible to calculate the limit of a function that tends to be infinite ∞ and identify some mathematical properties. In this way, it is also possible to internalize not just what has happened but also what may have happened.

3. WALLIS FORMULA AND STIRLING PROXY - JACEK CICHON

Let's consider the following integral (hence the area below the function between two points):



$$I(n) = \int_0^{\pi} \sin^n(x)\, dx$$

This is the Wallis integral, we can proceed by steps through the integration by parts

$$I(0) = \int_0^{\pi} \sin^0(x)\, dx = \int_0^{\pi} 1\, dx \rightarrow 1\Big|_0^{\pi} = [0 + \pi] = \pi$$

$$I(1) = \int_0^{\pi} \sin^1(x)\, dx = [\cos(0) - \cos(\pi)] = 2$$

And then:

$$I(2) = \int_0^{\pi} \sin^2(x)\, dx = \; = \int_0^{\pi} \frac{1-\cos(2x)}{2}\, dx = \frac{1}{2}\int_0^{\pi} 1-\cos(2x)\, dx = \frac{1}{2}\left(\left[x - \frac{\sin(2x)}{x}\right]_0^{\pi}\right) \rightarrow$$

$$\frac{1}{2}\left(\left[\pi - \frac{\sin(2\pi)}{\pi} + 0 - \frac{\sin(0)}{0}\right]\right) = \frac{1}{2}([\pi - 0 + 0 - 0]) = \frac{\pi}{2}$$

At this stage, we can calculate the integral for all the even values I(2n):

$$I(n>2) = \int_0^{\pi} \sin^n(x)\, dx = \int_0^{\pi} \sin^2(x) * \sin^{n-2}(x)\, dx = \int_0^{\pi} (1-\cos^2(x)) * \sin^{n-2}(x)\, dx =$$

$$\int_0^{\pi} \sin^{n-2}(x)\; dx - \int_0^{\pi} \sin^{n-2}(x)\cos^2(x)\, dx =$$

The second element can be integrated by the integration by parts:

- $f(x) = \cos(x) \rightarrow f'(x) = \sin(x)$
- $g'(x) = \sin^{n-2}(x)\cos(x) \rightarrow g(x) = \frac{1}{n-1}\sin^{n-1}(x)$ [5]:

Hence:

$$\int_0^{\pi} (\sin^{n-2}(x)\cos(x)) * \cos(x)\, dx$$
$$= \left[\frac{1}{n-1}\sin^{n-1}(x) * \cos(x)\right]_0^{\pi} + \frac{1}{n-1} * \int_0^{\pi} (\sin^{n-1}(x)\sin(x))\, dx$$

Then substituting:

$$I(n>2) = \int_0^{\pi} \sin^n(x) = \int_0^{\pi} \sin^{n-2}(x)\; dx - \frac{1}{n-1} * \int_0^{\pi} \sin^n(x)$$

Hence:

$$\int_0^{\pi} \sin^{n-2}(x)\; dx = (1 + \frac{1}{n-1}) \int_0^{\pi} \sin^n(x)$$
$$\int_0^{\pi} \sin^{n-2}(x)\; dx = \frac{n}{n-1}\int_0^{\pi} \sin^n(x)$$

Which can be written in the following form: $W_n = \int_0^{\pi} \sin^n(x)$

$$W_{n-2} = \frac{n}{n-1} * W_n \;\rightarrow\; W_n = W_{n-2} * \frac{n-1}{n}$$

---

[5] $\int_a^b f(x) * g'(x)dx = f(x) * g(x) - \int_a^b f(x)' * g(x)dx$

The integral of the product of a function "u" and a derivate "v" can be calculated as the difference between:

- The product of the function "u" and the primitive of function "v"
- The Integral of the product of the derivate of "u" and the primitive of function "v"

In this case we have applied to g'(x) again the integration by parts:

- u = sin(x), the anti-derivative is: $\frac{1}{n-1}\sin^{n-1}$
- v = cos(x), the derivative is -sin(x)

So, now, we have a recurrence formula:

$$W_n = \frac{n-1}{n} * \frac{n-2-1}{n-2} W_{n-4} \rightarrow$$

$$W_n = \frac{n-1}{n} * \frac{n-2-1}{n-2} * \frac{n-4-1}{n-4} * \dots * \frac{n-(n-2)-1}{n-(n-2)} * W_2$$

As we have seen above:

$$W_2 = \int_0^{\pi} \sin^2(x)\, dx = \frac{\pi}{2}$$

Hence:

$$W_{n,EVEN} = \frac{n-1}{n} * \frac{n-2-1}{n-2} * \frac{n-4-1}{n-4} * \dots \frac{n-(n-4)-1}{n-(n-4)} * \frac{n-(n-2)-1}{n-(n-2)} * \frac{\pi}{2}$$

We can immediately note that all the numerators are odd numbers, whilst the denominator are even:

$$W_{n,EVEN} = \frac{n-1}{n} * \frac{n-2-1}{n-2} * \dots \frac{5}{4} * \frac{3}{2} * \frac{\pi}{2}$$

Whilst, all the W_n are associated with only even numbers.

If we complete the series with the odd numbers:

$$W_{n,ODD} = \frac{n-1}{n} * \frac{n-2-1}{n-2} * \frac{n-4-1}{n-4} * \dots \frac{n-(n-4)-1}{n-(n-4)} * \frac{n-(n-2)-1}{n-(n-2)} * W_1$$

$$W_{n,ODD} = \frac{n-1}{n} * \frac{n-2-1}{n-2} * \frac{n-4-1}{n-4} * \dots \frac{n-(n-4)-1}{n-(n-4)} * \frac{n-(n-2)-1}{n-(n-2)} * 2$$

And again:

$$W_{n,ODD} = \frac{n-1}{n} * \frac{n-2-1}{n-2} * \dots \frac{6}{5} * \frac{4}{3} * 2$$

At this stage, we can note that:

$$W_{n+1} < W_n < W_{n-1}$$

The larger the power, the smaller the area underneath the curve, hence:

$$\frac{W_{n+1}}{W_{n+1}} < \frac{W_n}{W_{n+1}} < \frac{W_{n-1}}{W_{n+1}} \rightarrow 1 < \frac{W_n}{W_{n+1}} < \frac{W_{n-1}}{W_{n+1}}$$

Which, written in the form of a limit is, thanks to the squeeze theorem:[6]

$$\lim_{n \to \infty} \frac{W_n}{W_{n+1}} = 1$$

[6] Again, we do not provide here proof of such theorem.

Hence, if we calculate this limit:

$$\lim_{n\to\infty}\frac{W_{n,ODD}}{W_{n,EVEN}} = \lim_{n\to\infty}\frac{\frac{n-1}{n}*\frac{n-2-1}{n-2}*\ldots\frac{6}{5}*\frac{4}{3}*2}{\frac{n-1}{n}*\frac{n-2-1}{n-2}*\ldots\frac{5}{4}*\frac{3}{2}*\frac{\pi}{2}} \rightarrow$$

$$\lim_{n\to\infty}\frac{(2k)^2}{(2k-1)^2}*..*\frac{6^2}{5^2}*\frac{4^2}{3^2}*2^2*\frac{2}{\pi}=1$$

Hence:

$$\frac{(2k)^2}{(2k-1)^2}*..*\frac{6^2}{5^2}*\frac{4^2}{3^2}*2^2=\frac{\pi}{2}\rightarrow\frac{\pi}{2}=\frac{2}{1}*\frac{2}{3}*\frac{4}{3}*\frac{4}{5}*\frac{6}{5}*\ldots$$

Which is the Wallis product: the ratio between the products of the squared even numbers over the squared odd numbers

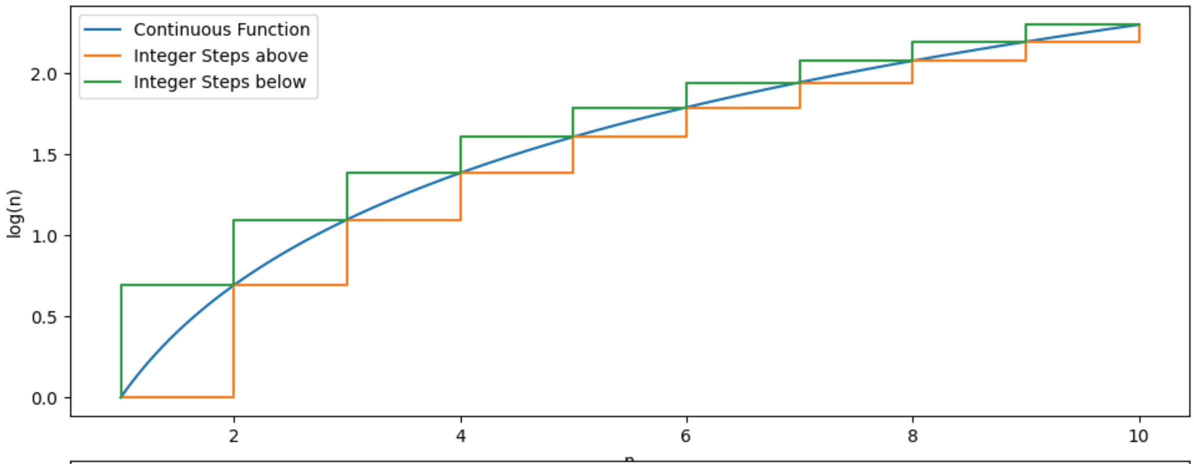
4. STIRLING FORMULA[7]

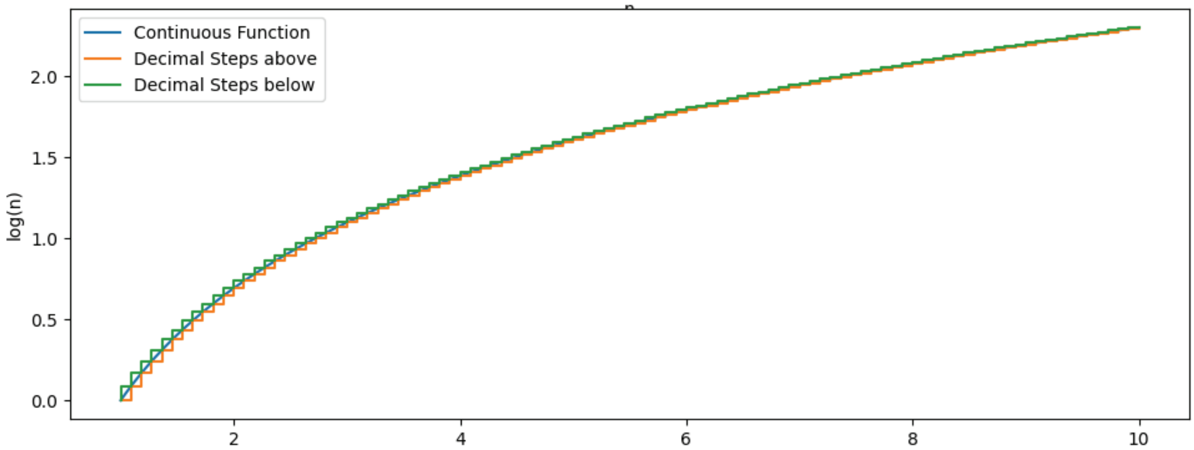
The factorial is a product of numbers, taking the logarithm of this measure we can represent it as a sum of integers.

$$n! \rightarrow \ln(n!) = \ln(1*2*3*\ldots*n) = \ln(1)+\ln(2)+..+\ln(n)$$

Actually, with factorial, we are not considering the continuous function, but just the discrete (hence a step function, like the one which is shown below).

[7] https://cs.pwr.edu.pl/cichon/Math/StirlingApp.pdf
https://www.matematicamente.it/forum/una-dimostrazione-elementare-della-formula-di-stirling-t107369.html





Considering the area below the curves (thereby the area below each step)

- The shorter the step, the larger the similarity between the steps functions;
- The shorter the step, the better their proxy to the continuous function.

The Integral is the function which allows to evaluation of the area underneath a curve, hence, through the integration by parts:

$$\int \ln(x)\,dx = x * \ln(x) - x = x * (\ln(x) - 1) = x * (\ln(x) - \ln(e)) = x * \ln(\frac{x}{e})$$

Hence:

$$\int_1^n \ln(x)\,dx = \left[x * \ln\left(\frac{x}{e}\right)\right]_1^n = n * \ln\left(\frac{n}{e}\right) - \ln\left(\frac{1}{e}\right) = \ln\left(\frac{n}{e}\right)^n + \ln(e) = \ln\frac{n^n}{e^{n-1}}$$

And:

$$\int_1^{n+1} \ln(x)\,dx = \cdots = \ln\frac{n^{n+1}}{e^n}$$

These two results represent the area underneath the steps functions (lower and upper), hence:

$$\ln\frac{n^n}{e^{n-1}} < n! < \ln\frac{n^{n+1}}{e^n}$$

This implies that "n!" is going to be a function of $\frac{n^n}{e^n}$ since:

$$\ln\frac{n^n}{e^n} * e < n! < \ln\frac{n^n}{e^n} * (n+1) * e \rightarrow n! = \alpha(n)\left(\frac{n}{e}\right)^n \rightarrow \alpha(n) = \frac{n!}{\left(\frac{n}{e}\right)^n}$$

Now, we have to prove that $\alpha(n) = \sqrt{2\pi n}$

Consider: $\alpha(n) \rightarrow b(n) = \ln(a(n)) = \ln\left(\frac{n!}{\sqrt{2n}\left(\frac{n}{e}\right)^n}\right)$

Hence:

$$b(n) - b(n+1) = 0.5(n+1) * \ln\left(\frac{n+1}{n}\right) - 1$$

At this stage, the Taylor expansion of:

$$\ln(1+x) = \sum\frac{x^{2k+1}}{2(k+1)} - \sum\frac{x^{2k}}{2(k)}$$

$$-\ln(1-x) = \sum\frac{x^{2k+1}}{2(k+1)} + \sum\frac{x^{2k}}{2(k)}$$

So that, their difference shows just the "odd" elements

Trying to apply such reasoning to our case

$$\ln\left(\frac{n+1}{n}\right) = \ln\left(\frac{1+x}{1-x}\right) \rightarrow (1+x) * n = (n+1)(1-x) \rightarrow \cdots \rightarrow x = \frac{1}{2n+1}$$

$$\mathrm{b}(x) - b(x+1) = \sum_{k=1}^{\infty}\frac{1}{2k+1} * \left(\frac{1}{2n+1}\right)^{2k} = \cdots = \frac{1}{4 * n * (n+1)}$$

Making the sum of all the components: b(n) – b(n+1)) + (b(n+1) – b(n+2)) + …

$$\mathrm{b}(1) - b(x) < \frac{1}{4}\sum_{m=1}^{n-1}\frac{1}{m(m+1)}$$

Hence

$$b(x) > b(1) - \frac{1}{4}$$

In such case:

$$\ln\big(b(n)\big) = \ln\left(\frac{n!}{\sqrt{2n}\left(\frac{n}{e}\right)^n}\right)$$

Its limit can be evaluated through Wallis formula which can be written as:

$$\lim \frac{2^{4n}(n!)^4}{(2n!)^4 * (2n+1)} = \frac{\pi}{2} \rightarrow$$

$$\lim \frac{2^{4n} * C^4 * (2n)^2 * \left(\frac{n}{e}\right)^{4n}}{C^2 * 4n * \left(\frac{2n}{e}\right)^{2n} * (2n+1)} = \frac{\pi}{2} \rightarrow C^2 * \lim \frac{2^{4n} * 4n^2 * n^{4n}}{4n * (2n+1)(2n)^{4n}} = \frac{\pi}{2}$$

$$C^2 * \lim \frac{n^2}{n * (2n+1)} = \frac{\pi}{2}$$

So $C = \sqrt{\pi}$

5. CONCLUSION:

HHaving applied the reductionism principles is not the easiest way to proceed since, although the it is possible to reach a conclusion, there may be some further internal operations or theorems which should not be given for granted (the explanation of which is not straightforward).

If, from one side, this does not impinge the validity of the results, on the other, it simply moves forward the efforts to fully explain a specific formula since the burden of proof is just re-located backwards to some other theorems.

6. REFERENCES:

Hald, Anders. A history of probability and statistics and their applications before 1750. John Wiley & Sons, 2005.
https://www.mdthinducollege.org/ebooks/statistics/A_History_of_Probability_and_Statistics.pdf

Jachek Cickon, Stirling Approximation formula (undated)
https://cs.pwr.edu.pl/cichon/Math/StirlingApp.pdf

Wigner, Eugene P. "The unreasonable effectiveness of mathematics in the natural sciences." Mathematics and science. 1990. 291-306.