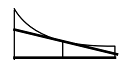
How did someone think up that inequality?

After all, if “ A > B ” is true so is “ A > B+ß ” for all sufficiently tiny ß > 0 as well as all ß < 0 ; of these infinitely many inequalities how did the one actually proved get chosen? In some instances the choice seems artificial, as if the proof had been devised first and then the result presented as a puzzle: “ *Find the proof.* ” This may explain why …

Many people, not just students, find inequalities too troublesome, and avoid them, leaving rewarding careers open to students willing to rise to the challenge.

Our first few inequalities provide relatively neat estimates for quantities that cannot be computed both exactly and neatly. These first few arise from the observation that the graph of a *convex* function $f(t)$ (curved like  because $d^2f(t)/dt^2 > 0$) lies below its secants but above its tangents. Therefore the area under the graph between x and y satisfies inequalities like

 $= (y-x)\cdot(f(x)+f(y))/2 > \int_x^y f(t)dt > (y-x)\cdot f((x+y)/2) =$  .

Summing appropriate inequalities delivers the desired final results. But first something easier:

0. Bernoulli’s Inequality:

Prove that if real $x > -1$, and if real $p \le 0$ or $p \ge 1$, then $(1 + px) \le (1 + x)^p$. (Graph it!)

Proof: This ancient inequality dates from the early years of the Calculus. Geometrically it says that the graph of $(1 + x)^p$ is convex (U - shaped) and therefore lies above its every tangent, particularly the tangent drawn through the point on the curve where $x = 0$. Our analytic proof will start from the derivative $d(1 + x)^p/dx = p(1 + x)^{p-1}$. Integration yields

$$\int_0^X p\left((1 + x)^{p-1} - 1\right) dx = (1 + X)^p - 1 - pX .$$

For all x between 0 and X the integrand $p((1+x)^{p-1} - 1)$ has the same sign as X because

- if $X > 0$ and $p > 1$ then $p((1+x)^{p-1} - 1) > 0$;
- if $X > 0$ and $p < 0$ then $p((1+x)^{p-1} - 1) > 0$;
- if $-1 < X < 0$ and $p > 1$ then $p((1+x)^{p-1} - 1) < 0$;
- if $-1 < X < 0$ and $p < 0$ then $p((1+x)^{p-1} - 1) < 0$.

Therefore the integral is nonnegative, which confirms Bernoulli’s Inequality. This inequality gets reversed if $0 \le p \le 1$; can you see why? (DRAW GRAPHS !)

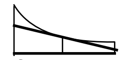
1. Harmonic Numbers

The k^{th} Harmonic Number is $H_k := 1/1 + 1/2 + 1/3 + \ldots + 1/(k-1) + 1/k$ for integers $k > 0$. Prove $\ln(k + 1/2) - \ln(m + 1/2) \ge H_k - H_m \ge \ln(k) + 1/(2k) - \ln(m) - 1/(2m)$ whenever $k \ge m > 0$. (When $m = 2$ these two inequalities bracket H_k within 1% for all $k \ge 2$.)

Proof: H_k will be estimated from the integral $\int_x^{(x+1)} \frac{1}{t} dt = \ln(x + 1) - \ln(x)$ for $x > 0$.

For instance, $1/x > 1/t > 1/(x+1)$ inside the integral, so $1/x > \ln(x+1) - \ln(x) > 1/(x+1)$; and then $\ln(k+1) < H_k < 1 + \ln(k)$ would follow by summing appropriate inequalities. A much better

estimate comes from the observation that the graph of $y = 1/t$ is *convex* (curved like 
because $d^2y/dt^2 = 2/t^3 > 0$) and thus lies below its secants but above its tangents. Consequently areas under the curve satisfy

 $= (1/x + 1/(x+1))/2 > \ln(x+1) - \ln(x) > 1/(x + 1/2) =$  .

Summing appropriate inequalities (can you see which?) now establishes for $k \geq m > 0$ that

$$\ln(k + 1/2) – \ln(m + 1/2) \geq H_k – H_m \geq \ln(k) + 1/(2k) – \ln(m) – 1/(2m) .$$

2. Sums of Reciprocal Squares

For any integer $k > 0$ we seek close estimates for $S_k := 1/1 + 1/4 + 1/9 +1/16 + \dots + 1/k^2$. An easy estimate starts from the easily proved (how?) *Telescoping* formula …

$$1/(m(m+1)) + 1/((m+1)(m+2)) + \dots + 1/((k–1)k) = 1/m – 1/k .$$

Use this to prove that $S_k < 2 – 1/k$ for $k > 1$. Next prove the better but messier estimate …

$$1/m – 1/(2m^2) – 1/k + 1/(2k^2) < S_k – S_m < 1/(m + 1/2) – 1/(k + 1/2) \quad \text{whenever} \quad k > m > 0 .$$

(When $m = 2$ these two inequalities bracket S_k well within 2% for all $k \geq 2$.)

Proofs: The Telescoping formula above is proved by induction on $k–m = 0, 1, 2, 3, \dots$ for any integer $m > 0$ from the observation that $1/k – 1/(k+1) = 1/(k(k+1))$, and then the formula is applied with $m = 1$ and the observation that $1/k^2 < 1/((k–1)k)$ to infer $S_k – S_1 < 1 – 1/k$. To prove the "messier estimate", use the integral $\int_x^{(x+1)} t^{-2}dt = \frac{1}{x} - \frac{1}{x+1}$ for $x > 0$. Since the graph of $1/t^2$ is convex it lies below its secants but above its tangents (see Harmonic Numbers above); consequently $(1/x^2 + 1/(x+1)^2)/2 > 1/x – 1/(x+1) > 1/(x + 1/2)^2$.

(These two inequalities can be proved by algebraic means alone with no appeal to Calculus; can you see how?)

Summing appropriate inequalities (can you see which?) now establishes for $k > m > 0$ that $1/m – 1/(2m^2) – 1/k + 1/(2k^2) < S_k – S_m < 1/(m + 1/2) – 1/(k + 1/2)$, as was required.

This provides an estimate of sorts for a value of Riemann's Zeta function $\zeta(2) = S_\infty = \pi^2/6$ in terms of S_m for large m , namely $S_m + 1/m – 1/(2m^2) < \pi^2/6 < S_m + 1/(m + 1/2)$ within an interval narrower than $1/(4m^3)$.

3. Sums of Reciprocals of Square Roots

For any integer $k > 0$ we seek close estimates for $R_k := 1/\sqrt{1} + 1/\sqrt{2} + 1/\sqrt{3} + \dots + 1/\sqrt{k}$. An easy estimate starts from a Telescoping formula …

$$1/(\sqrt{m} + \sqrt{(m+1)}) + 1/(\sqrt{(m+1)} + \sqrt{(m+2)}) + \dots + 1/(\sqrt{(k–1)} + \sqrt{k}) = \sqrt{k} – \sqrt{m} .$$

Prove this and then prove $R_k > 2\sqrt{(k+1)} – 2$. Next prove the better but messier estimate …

$$2\sqrt{k} + 1/(2\sqrt{k}) – 2\sqrt{m} – 1/(2\sqrt{m}) < R_k – R_m < 2\sqrt{(k + 1/2)} – 2\sqrt{(m + 1/2)} \quad \text{whenever} \quad k > m > 0 .$$

(When $m = 2$ these two inequalities bracket R_k well within 1% for all $k \geq 2$. In particular,
$R_k \geq 2\sqrt{k} + 1/(2\sqrt{k}) + 1 – 7/\sqrt{8} > 2\sqrt{(k+1)} – 2$; can you prove the last inequality?)

Proofs: The Telescoping formula here is proved by induction again and provides a proof that

$$R_k > 2/(\sqrt{1} + \sqrt{2}) + 2/(\sqrt{2} + \sqrt{3}) + 2/(\sqrt{3} + \sqrt{4}) + \dots + 2/(\sqrt{k} + \sqrt{(k+1)}) = 2\sqrt{(k+1)} – 2 .$$

Better estimates for R_k come from the integral

$$\int_x^{(x+1)} t^{-1/2}dt = 2\sqrt{(x+1)} - 2\sqrt{x} \quad \text{for } x > 0 .$$

Since the graph of $1/\sqrt{t}$ is convex it lies below its secants but above its tangents as before, so

$$(1/\sqrt{x} + 1/\sqrt{(x+1)})/2 > 2\sqrt{(x+1)} - 2\sqrt{x} > 1/\sqrt{(x + 1/2)} .$$

(These two inequalities can be proved by algebraic means alone with no appeal to Calculus; can you see how?)

Summing appropriate inequalities establishes for $k > m > 0$ the “messier estimate” required:

$$2\sqrt{k} + 1/(2\sqrt{k}) - 2\sqrt{m} - 1/(2\sqrt{m}) < R_k - R_m < 2\sqrt{(k + 1/2)} - 2\sqrt{(m + 1/2)} .$$

Is this estimate really better than can be obtained from the Telescoping formula? For $k \ge m = 2$ we find that $R_k \ge 2\sqrt{k} + 1/(2\sqrt{k}) + 1 - 7/\sqrt{8} > 2\sqrt{(k+1)} - 2$, though the last inequality is not obvious. It can be inferred easily from an equivalent inequality $2/(\sqrt{k} + \sqrt{(k+1)}) - 1/(2\sqrt{k}) < 1/(2\sqrt{k}) \le 1/\sqrt{8} < 3 - 7/\sqrt{8}$.

A pattern is emerging for these sums of series. To see how far this pattern can go look up the *Euler-Maclaurin Sum Formula* in *Advanced Calculus* texts or old *Numerical Analysis* texts. In these texts repose several centuries’ lore about rapid approximate computations of functions whose exact computation would be intolerably onerous. One more example follows:

4. James Stirling’s Approximation: $n! \approx \sqrt{2\pi \cdot n} \cdot (n/e)^n$.

It was published in 1730 . The formula’s *relative* (not *absolute*) error approaches zero as n approaches $+\infty$. For example …

Table 1: Stirling’s Approximation

n	n**!**	$\sqrt{2\pi \cdot n} \cdot (n/e)^n$	Rel. error
10	3,628,800	$3.60 \cdot 10^6$	0.8 %
20	$2.433 \cdot 10^{18}$	$2.423 \cdot 10^{18}$	0.4 %
40	$8.159 \cdot 10^{47}$	$8.142 \cdot 10^{47}$	0.2 %
80	$7.157 \cdot 10^{118}$	$7.149 \cdot 10^{118}$	0.1 %
160	$4.715 \cdot 10^{284}$	$4.712 \cdot 10^{284}$	0.05 %

A much better approximation can be obtained from the (nonconvergent!) *Asymptotic Series*

$$n! \approx \sqrt{2\pi \cdot n} \cdot (n/e)^n \cdot \exp\big(1/(12 \cdot n) - 1/(360 \cdot n^3) + 1/(1260 \cdot n^5) - 1/(1680 \cdot n^7) + \dots \big) \quad \text{for large } n ,$$

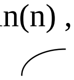
but it lies far beyond the scope of this course. Instead prove the weaker approximations …

$$(n+1/2) \cdot \ln(n+1/2) - n - (3/2) \cdot \ln(3/2) + 1 < \ln(n!) < (n+1/2) \cdot \ln(n) - n + 2 - (3/2) \cdot \ln(2) .$$

(These inequalities bracket n**!** within 7% because the upper bound exceeds the lower by

$$1 - (3/2) \cdot \ln(4/3) - (n+1/2) \cdot \ln((n+1/2)/n) = (1/2) \cdot \ln(1-z)/z + 0.568477\dots < 0.0685 ,$$

where $z := 1/(2n+1)$ and $\ln(1-z)/z = -1 - z/2 - z^2/3 - z^3/4 - z^4/5 - \dots$.)

Proof: We shall exploit the integral $\int \ln(x) \cdot dx = x \cdot \ln(x) - x$ to estimate upper and lower bounds for the finite series $\ln(n!) = \sum_{1 \le k \le n} \ln(k) = \ln(2) + \ln(3) + \ln(4) + \dots + \ln(n-1) + \ln(n)$, when $n > 1$, as was done before except that the graph of $\ln(x)$ is *concave* (curved like ) now because $\ln(x)'' = -1/x^2 < 0$, so the graph lies *above* its secants but *below* its tangents.

Consequently

$$(\ln(x) + \ln(x+1))/2 \; < \; \int_x^{(x+1)} \ln(t)\,dt \; = \; (x+1)\cdot\ln(x+1) - 1 - x\cdot\ln(x) \; < \; \ln(x+1/2) \,.$$

As before, summing appropriate inequalities (you should figure out which) now implies

$$(n+1/2)\cdot\ln(n+1/2) - n - (3/2)\cdot\ln(3/2) + 1 \; < \ln(n!) < \; (n+1/2)\cdot\ln(n) - n + 2 - (3/2)\cdot\ln(2) \;,$$

as claimed. The upper bound exceeds the lower by

$$1 - (3/2)\cdot\ln(4/3) - (n+1/2)\cdot\ln((n+1/2)/n) \; = \; (1/2)\cdot\ln(1-z)/z + 0.568477\ldots \; < \; 0.0685 \,,$$

where $z := 1/(2n+1)$ and $\ln(1-z)/z = -1 - z/2 - z^2/3 - z^3/4 - z^4/5 - \ldots < -1$. Consequently

$$0.89 < 0.962\ldots - 0.0685 < \; Ç(n) := \ln(n!) - (n+1/2)\cdot\ln(n) + n \; < \; 2 - (3/2)\cdot\ln(2) = 0.962\ldots \,.$$

This $Ç(n)$ is a decreasing function of n because, after some algebra,

$$Ç(n+1) - Ç(n) = \; 1 + (1/2)\cdot\ln((1-z)/(1+z))/z \; = \; -z^2/3 - z^4/5 - z^6/7 - \ldots \; < 0 \,.$$

Therefore, as n increases towards infinity, $Ç(n)$ decreases towards a limit $Ç > 0.89$. Although Stirling did not know it at first, this constant $Ç$ turns out to be $\ln(\sqrt{2\pi}) = 0.919\ldots$, as shall be proved after the next problem. For now we conclude, for some constant $Ç$ between 0.962 and 0.89 , that $\ln(n!) - (n+1/2)\cdot\ln(n) + n - Ç$ approaches zero or, equivalently, that $n!/(\,e^{Ç}\cdot\sqrt{n}\cdot(n/e)^n\,)$ decreases towards 1 in the limit as n increases towards infinity.

5. For integers $m \ge 0$ set $J_m := \int_0^{\pi/2} (\sin x)^m dx$ and prove $J_m = (1 - 1/m)\cdot J_{m-2}$ for $m \ge 2$. Then confirm the formulas $J_{2k+1} = (2^k\cdot k!)^2/(2k+1)!$ and $J_{2k} = (2k)!\cdot(\pi/2)/(2^k\cdot k!)^2$ for $k \ge 0$, and prove that $J_m/J_{m-1} \to 1$ as $m \to +\infty$.

Proofs: For $m \ge 2$ integration by parts yields

$$J_m = \int_0^{\pi/2} (\sin x)^m dx \; = \; -\int_0^{\pi/2} (\sin x)^{m-1} d\cos x \; = \; \int_0^{\pi/2} (\cos x)\, d(\sin x)^{m-1} \quad (\text{ using } \textit{I-by-P}\,)$$

$$= (m-1)\int_0^{\pi/2} (\cos x)^2 (\sin x)^{m-2} dx \; = \; (m-1)\int_0^{\pi/2} (1 - (\sin x)^2)\,(\sin x)^{m-2} dx \;,$$

from which follows the recurrence $J_m = (m-1)\cdot(J_{m-2} - J_m) = (1 - 1/m)\cdot J_{m-2}$ starting from

$$J_1 = \int_0^{\pi/2} (\sin x)^1 dx \; = \; 1 \quad \text{and} \quad J_0 = \int_0^{\pi/2} (\sin x)^0 dx \; = \; \pi/2 \,.$$

Now induction on $k = 0, 1, 2, 3, \ldots$ in turn provides confirmation for the formulas

$$J_{2k+1} = (2^k\cdot k!)^2/(2k+1)! \quad \text{and} \quad J_{2k} = \; (2k)!\cdot(\pi/2)/(2^k\cdot k!)^2 \,.$$

Moreover, because $0 < \sin x < 1$ inside the range of integration, $0 < J_m < J_{m-1}$. Consequently $1 > J_m/J_{m-1} = (1 - 1/m)\cdot J_{m-2}/J_{m-1} > (1 - 1/m) \to 1$ and therefore $J_m/J_{m-1} \to 1$ as $m \to +\infty$, and so does $(\pi/2)\cdot J_{2k+1}/J_{2k} \; = \; (2^k\cdot k!)^4/((2k+1)!\cdot(2k)!) \; \to \; \pi/2$ as $k \to +\infty$. That quotient of factorials *etc*. provides an estimate for $\pi/2$ found first by John Wallis (who died in 1730), and leads as follows to the value for $Ç$ defined in the previous problem.

Replace each factorial in that quotient by its Stirling approximation $n! \approx e^{Ç}\cdot\sqrt{n}\cdot(n/e)^n$ and let $k \to +\infty$. We find that Stirling's approximation to $(2^k\cdot k!)^4/((2k+1)!\cdot(2k)!)$ simplifies, after a lot of algebra, to $e^{2Ç+1}\cdot 2^{-2}\cdot(1 + 1/(2k))^{-2k-3/2} \to e^{2Ç}/4$ as $k \to +\infty$ since $(1 + 1/(2k))^{2k} \to e$. This implies that $e^{2Ç}/4 = \pi/2$, whence $e^{Ç} = \sqrt{2\pi}$, completing the vindication of

$$\text{Stirling's Approximation} \quad n! \; \approx \; \sqrt{2\pi\cdot n}\cdot(n/e)^n \,.$$

6. Arithmetic *vs*. Geometric Means

Given collections of positive variables x_j and positive *weights* w_j, where we restrict subscript j to some finite set, let

$$w := \sum_j w_j \,, \qquad A := (\sum_j w_j \cdot x_j)/w \quad \text{and} \qquad G := \left(\prod_j x_j^{\,w_j} \right)^{1/w} .$$

Here A is the *Weighted Arithmean Mean* (Average) of the numbers x_j, and G is their *Weighted Geometric Mean*. Prove that $A \ge G$ with equality just when all the x_j 's have the same positive value.

Short proof: First simplify the notation by defining *fractional weights* $f_j := w_j/w > 0$ so that

$$\sum_j f_j = 1 \,, \qquad A := \sum_j f_j \cdot x_j \quad \text{and} \qquad G := \prod_j x_j^{\,f_j} .$$

Next observe for any $x > 0$ that $0 \le \int_x^G \left(\frac{1}{t} - \frac{1}{G} \right) dt \;=\; \ln(G) - \ln(x) - 1 + x/G$ because, so long as the integrand's t lies strictly between x and G, the signs of $G–x$ and of $1/t - 1/G$ must be the same. Of course " $0 \le \dots$ " becomes " $0 = \dots$ " just when $x = G$. Now replace x by x_j, multiply by f_j, and sum over j to deduce that $0 \le 0 - 1 + A/G$ as was claimed.

Another proof uses Jensen's Inequality; see the notes about it on the class web page.