THE PROBABILITY INTEGRAL

'On a donc fait un hypothèse, et cette hypothèse a été appelée loi des erreurs. Elle ne s'obtient pas des déductions rigoreuses …"Tout le monde y croit cependent," me disait un jour M Lippman, "car les expérimenteurs s'imaginent que c'est un théorems de mathématiques, et les mathématiciens que c'est un fait expérimental", H Poincaré, *Calcul des Probabilités*, Paris: Gauthier-Villars 1896 and 1912.

We say that Z has a standard normal distribution if it has the probability density function

$$f_Z(z) = \phi(z)$$

where $\phi(z)$ is the function

$$\phi(z) = \frac{1}{\sqrt{2\pi}} \exp(-\tfrac{1}{2}z^2).$$

According to Gnedenko, §22, the integral $\int_{-\infty}^{+\infty} \phi(z)\, dz$ is called the Poisson integral. Although this function is clearly non-negative, it is by no means clear that it integrates to unity. There are a number of methods of showing that

$$I = \int_0^\infty \exp(-\tfrac{1}{2}z^2)\, dz = \sqrt{\frac{\pi}{2}}$$

none of which is obvious.

1. De Moivre showed (see A De Moivre, *Approximatio ad Summam Terminorum Binomii* $\overline{a+b}\backslash^n$ *in Seriem expansi* 1733, reprinted in R C Archibald, A rare pamphlet of Moivre and some or his discoveries, *Isis* **8** (1926), 671–683; translated with some additions in A De Moivre, *The Doctrine of Chances* (2nd edn), London: H Woodfall 1738, reprinted London: Cass 1967, A De Moivre, *The Doctrine of Chances* (3rd edn), London: A Millar 1756, reprinted New York, NY: Chelsea 1967 with a biographical article from *Scripta Mathematica* **2**(4) (1934), 316–333 by H M Walker) that if $n = 2m$ and

   $$b(x) = \binom{2m}{x}\left(\frac{1}{2}\right)^x\left(\frac{1}{2}\right)^{2m-x} = \binom{2m}{x}2^{-2m}$$

   then

   $$b(x) \sim \frac{0.7976}{\sqrt{n}}$$

   and the exact value of the constant was shown by James Stirling to be $\sqrt{2/\pi}$ (see A Hald, *A History of the Theory of Probability and Statistics and Their Application before 1750*, New York, NY, etc: Wiley 1990, §24.4). He then went on to show that

   $$b(m+l) \sim b(m)\exp(-2l^2/n) \sim \sqrt{\frac{2}{\pi n}}\exp(-2l^2/n).$$

   With σ^2 taking its binomial value $pqn = n/4$ this is of the form

   $$b(m+l) \sim \frac{1}{\sqrt{2\pi\sigma^2}}\exp(-\tfrac{1}{2}l^2/\sigma^2) = \sigma^{-1}\phi(l/\sigma)$$

   but De Moivre did not use the concept of variance and did not express it that way. Using the binomial theorem it could be concluded that

   $$\begin{aligned} 1 = \sum_{l=-m}^{+m} b(m+l) &\sim \int_{-m}^{+m} \sigma^{-1}\phi(l/\sigma)\, dl \\ &\sim \int_{-\infty}^{+\infty} \phi(x)\, dx. \end{aligned}$$

   It is not entirely simple to justify the limiting processes. De Moivre did not in fact make any remark about the integral as such.

2. The first method by which the integral was explicitly calculated appears to have been given by P S Laplace in his 1774 paper Mémoire sur la probabilité des causes par les évenments. An extract from this paper reads as follows: "... From this we can easily conclude

$$E = \frac{(p+1)\cdots(p+q+1)}{1\cdot 2\cdot 3\cdot\cdots\cdot q}.\frac{p^p q^q}{(p+q)^{p+q}}\int 2\,dz\,\cdot\exp\left(-\frac{(p+q)^3}{2pq}zz\right).$$

Let $-[(p+q)^3/2pq]zz = \ln\mu$, and we will have

$$\int 2\,dz\,\cdot\exp\left(-\frac{(p+q)^3}{2pq}zz\right) = -\frac{\sqrt{2qp}}{(p+q)^2}\int\frac{d\mu}{\sqrt{-\ln\mu}}.$$

The number μ can here have any value between 0 and 1, and, supposing the integral begins at $\mu = 1$, we need its value at $\mu = 0$. This may be determined using the following theorem (see M. Euler's Calcul intégral). Supposing the integral goes from $\mu = 0$ to $\mu = 1$ we have [1]

$$\int\frac{\mu^n\,d\mu}{\sqrt{(1-\mu^{2i})}}\cdot\int\frac{\mu^{n+i}\,d\mu}{\sqrt{(1-\mu^{2i})}} = \frac{1}{i(n+1)}\cdot\frac{\pi}{2},$$

whatever be n and i. Supposing $n = 0$ and i is infinitely small, we will have $(1-\mu^{2i})/(2i) = -\ln\mu$, because the numerator and the denominator oof this quantity become zero when $i = 0$, and if we differentiate them both, regarding i alone as variable, we will have $(1-\mu^{2i})/(2i) = \ln\mu$, therefore $1-\mu^{2i} = -2i\ln\mu$. Under these conditions we will thus have

$$\int\frac{\mu^n\,d\mu}{\sqrt{(1-\mu^{2i})}}\cdot\int\frac{\mu^{n+i}\,d\mu}{\sqrt{(1-\mu^{2i})}} = \int\frac{d\mu}{\sqrt{2i}\sqrt{-\ln\mu}}\int\frac{d\mu}{\sqrt{2i}\sqrt{-\ln\mu}} = \frac{1}{i}\frac{\pi}{2};$$

Therefore

$$\int\frac{d\mu}{\sqrt{-\ln\mu}} = \sqrt{\pi}.$$

Thus

$$\int 2\,dz\,\cdot\exp\left(-\frac{(p+q)^3}{2pq}zz\right) = \frac{\sqrt{pq}\sqrt{2\pi}}{(p+q)^{3/2}},$$

from which we obtain $E = 1$." Taking $p = q = 1/2$ in Laplace's result we get

$$\int_0^\infty \exp(-2z^2)\,dz = \sqrt{\pi/8}$$

and so on taking $x = 2z$

$$\int_0^\infty \exp\left(-\tfrac{1}{2}x^2\right)\,dz = \sqrt{\pi/2}$$

or in other words $I = \sqrt{\pi/2}$.

---

[1] Cf. the equation

$$\int_0^{\pi/2}\sin^\alpha dt = \frac{\sqrt{\pi}}{\alpha}\frac{\Gamma((1+\alpha)/2)}{\Gamma(\alpha/2)}$$

(R Courant, *Differential and Integral Calculus* (2 volumes), London and Glasgow: Blackie 1934–6, Volume II, Chapter IV, §6, page 338). We use the substitution $\mu^n = \cos\theta$ to reduce the integrals to

$$\int\frac{1}{i}\cos^{(n-i+1)/i}\theta\,d\theta \qquad\text{and}\qquad \int\frac{1}{i}\cos^{(n+1)/i}\theta\,d\theta.$$

Essentially what it needs is that if

$$I_k = \int_0^{\pi/2}\cos^k\theta\,d\theta$$

then

$$I_k I_{k+1} = \frac{\pi}{2(k+1)},$$

which as $I_k = B((k+1)/2, 1/2)/2$ is easily seen to be equivalent to $\Gamma(1/2)^2 = \pi$.

3. The integral is commonly evaluated using a double integral. The first method based on a double integral depends on noting that

$$I = \int_0^\infty \exp(-\tfrac{1}{2}z^2)\,dz = \int_0^\infty \exp(-\tfrac{1}{2}(xy)^2)\,y\,dx$$

for any y (on setting $z = xy$). Putting z in place of y, it follows that for any z

$$I = \int_0^\infty \exp(-\tfrac{1}{2}(zx)^2)\,z\,dx$$

so that

$$I^2 = \left(\int_0^\infty \exp(-\tfrac{1}{2}z^2)\,dz\right)\left(\int_0^\infty \exp(-\tfrac{1}{2}(zx)^2)z\,dx\right) = \int_0^\infty \int_0^\infty \exp\{-\tfrac{1}{2}(x^2+1)z^2\}\,z\,dz\,dx.$$

Now set $(1+x^2)z^2 = 2t$ so that $z\,dz = dt/(1+x^2)$ to get

$$\begin{aligned} I^2 &= \int_0^\infty \int_0^\infty \exp(-t)\,\frac{dt}{(1+x^2)}\,dx = \left(\int_0^\infty \exp(-t)\,dt\right)\left(\int_0^\infty \frac{dx}{(1+x^2)}\right) \\ &= [-\exp(-t)]_0^\infty\ \left[\tan^{-1}x\right]_0^\infty = [1]\left[\tfrac{1}{2}\pi\right] \\ &= \frac{\pi}{2} \end{aligned}$$

and hence $I = \sqrt{\pi/2}$ so that the integral of ϕ from $-\infty$ to ∞ is 1, and hence ϕ *is* a probability density function. This method is apparently due to P.S. Laplace (1749–1827), *Théorie Analytiques des Probabilités*, §24, pages 94–95 in the first edition.; cf. I Todhunter, *A History of the Mathematical Theory of Probability from the time of Pascal to that of Laplace*, Cambridge and London: Macmillan 1865, reprinted New York, NY: Chelsea 1949, art. 899. See, e.g., G Valiron, *Cours d'Analyse Mathématique* (2 volumes), Paris: Masson 1947, Volume I, page 152.

4. The usual "double integral" method is based on defining I as above and noting that

$$I^2 = \left(\int_0^\infty \exp(-\tfrac{1}{2}x^2)\,dx\right)\left(\int_0^\infty \exp(-\tfrac{1}{2}y^2)\,dy\right) = \int_0^\infty \int_0^\infty \exp\{-\tfrac{1}{2}(x^2+y^2)\}\,dx\,dy.$$

We then change to polar co-ordinates (r,θ) in which $dx\,dy = r\,dr\,d\theta$, so that

$$\begin{aligned} I^2 &= \int_0^{\pi/2}\int_0^\infty \exp\left(-\tfrac{1}{2}r^2\right)\,r\,dr\,d\theta = \left(\int_0^{\pi/2} d\theta\right)\left(\int_0^\infty \exp\left(-\tfrac{1}{2}r^2\right)\,r\,dr\right) = [\theta]_0^{\pi/2}\left[-\exp\left(-\tfrac{1}{2}r^2\right)\right]_0^\infty \\ &= \frac{\pi}{2} \end{aligned}$$

from which it follows that $I = \sqrt{\pi/2}$ so that the integral of ϕ from $-\infty$ to ∞ is 1, and hence ϕ *is* a probability density function. This method is apparently due to Siméon Denis Poisson (1781–1840) and was popularized by Jacob Karl Franz Sturm (1803–1855)—see his *Cours d'Analyse de l'école polytechnique*, Paris: Mallet-Bachelier, Volume 2, pages 16–17 which reads as follows:

"466. L'integrale

$$A = \int_0^\infty e^{-x^2}dx$$

a été déterminée par M. Poisson à l'aide d'un procédé très-remarquable. Si l'on change x en y, on aura encore

$$A = \int_0^\infty e^{-y^2}dy,$$

et, par suite,

$$A^2 = \int_0^\infty e^{-x^2}dx.\int_0^\infty e^{-y^2}dy = \int_0^\infty\int_0^\infty e^{-x^2-y^2}dxdy.$$

Soient mantainent trois axes rectangulaire Ox, Oy, Oz et

$$y = 0, \quad z = e^{-x^2},$$

les equations d'une courbe située dans le plan $z\mathrm{O}x$. Si cette courbe tourne autour de l'axe $\mathrm{O}z$, elle engendrera une surface ayant pour équation

$$z = e^{-x^2-y^2},$$

et l'intégrale double

$$\int_0^\infty \int_0^\infty e^{-x^2-y^2} dxdy,$$

réprésentera le quart du volume compris entre le surface et le plan $x\mathrm{O}y$. On peut évaluer ce volume par le partageant en une infinité de tranches cylindriques dont $\mathrm{O}z$ soit l'axe commun. La tranche dont les surfaces extérieurs ont pour rayone r et $r + dr$ est égale à sa base $2\pi r dr$ multipliée par sa hauteur z on e^{-r^2}: on a donc

$$A^2 = \frac{1}{4}\int_0^\infty e^{-r^2} \times 2\pi r dr = \frac{1}{4}\pi$$

d'où

$$A = \frac{1}{2}\sqrt{\pi}.\text{''}$$

5. Another method comes from the fact that

$$\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin \pi z}$$

with $z = \frac{1}{2}$—see E T Whittaker and G N Watson, *A Course of Modern Analysis*, Cambridge University Press 1902, 1915, 1920 and 1927. §12.14, J C Burkill and H Burkill, *A Second Course in Mathematical Analysis*, Cambridge University Press 1970, §14.6, or E T Copson, *Theory of Functions of a Complex Variable*, Oxford: Clarendon Press 1935, §9.22.

6. Yet another method results from substituting $u = \exp(-\frac{1}{2}z^2)$, giving

$$I = \int_0^1 \frac{du}{\sqrt{-2\ln u}}.$$

Now note that, for $x > 0$,

$$\frac{x-1}{x} \leqslant \ln x \leqslant x - 1$$

which follows geometrically from the convexity of the logarithmic function, or can be easily established using calculus to show, for example, that $x - \ln x$ has smallest value 1.

For $0 < v < 1$ and for any positive integer n, write $v_n = v^{1/n}$, so that $\ln v = n \ln v_n$. From the above inequalities with $x = v_n$,

$$n\left(\frac{v_n - 1}{v_n}\right) \leqslant \ln v \leqslant n(v_n - 1)$$

from which

$$\frac{1}{\sqrt{2n}}\sqrt{\frac{v_n}{1-v_n}} \leqslant \frac{1}{\sqrt{-2\ln v}} \leqslant \frac{1}{\sqrt{2n}}\frac{1}{\sqrt{1-v_n}}.$$

Integrating these inequalities between 0 and 1, we obtain

$$J_n \leqslant I \leqslant K_n$$

where

$$J_n = \frac{1}{\sqrt{2n}}\int_0^1 \sqrt{\frac{v_n}{1-v_n}} dv \qquad \text{and} \qquad K_n = \frac{1}{\sqrt{2n}}\int_0^1 \frac{1}{1-v_n} dv.$$

Now substitute $v_n = \sin^2\phi$, i.e. $v = \sin^{2n}\phi$. Then

$$J_n = \sqrt{2n}\int_0^{\pi/2} \sin^{2n}\phi d\phi \qquad \text{and} \qquad K_n = \sqrt{2n}\int_0^{\pi/2} \sin^{2n-1}\phi d\phi.$$

It is thus clear (as integrals of powers of $\sin\phi$ must decrease with the power involved) that

$$0 \leqslant \sqrt{\frac{n}{n+1}} K_{n+1} \leqslant J_n \leqslant K_n.$$

Furthermore, by the usual reduction method

$$\frac{J_{n+1}}{J_n} = \frac{K_n}{K_{n+1}} = \frac{2n+1}{2\sqrt{n(n+1)}} > 1$$

so that

$$J_{n+1}K_{n+1} = J_nK_n = \cdots = J_1K_1 = \pi/2.$$

It follows that, as $n \to \infty$, K_n decreases and J_n increases to a common limit $\sqrt{\pi/2}$. It follows that as $J_n \leqslant I \leqslant K_n$, we have $I = \sqrt{\pi/2}$. This method can be found in N Gauthier, Note 72.22 Evaluating the probability integral, *Mathematical Gazette*, **72** (1988), 124–125, and D Desbrow, Note 74.28 Evaluating the probability integral, *Mathematical Gazette* **74** (1990), 169–170, but I am not sure whether it originated there.

7. It was long supposed that the integral could not be evaluated by the Cauchy method of residues, but it turns out that it can be (see, e.g., J C Burkill and H Burkill, *A Second Course in Mathematical Analysis*, Cambridge University Press 1970, Exercises 14(a), no. 15). This method depends on setting

$$f(z) = \exp\left(\pi i z^2\right) / \sin(\pi z)$$

which function has residue $1/\pi$ at $z = 0$. Then since

$$\sin\{\pi(z-1)\} = -\sin\{\pi z\}$$

and

$$z^2 = z(z-1) + z \qquad \text{and} \qquad (z-1)^2 = z(z-1) - z + 1$$

we see that

$$f(z) - f(z-1) = 2i\exp\{\pi i z(z-1)\}.$$

By integrating f round a parallelogram with vertices $\pm\frac{1}{2} \pm R\exp(\frac{1}{4}\pi i)$, where R is large, (putting $z = t\exp(i\pi/4) \pm\frac{1}{2}$ so $dz = dt\exp(i\pi/4)$ along long sides) we see that

$$\int_{-\infty}^{\infty} \exp(-\pi t^2)dt = 1.$$

This method was due to G. Pólya (1949). A previous method using contour integrals due to L.J. Mordell (1920) "contains [the probability integral] as a special case, [but] the methods used by Mordell are too complicated and it is not really worthwhile applying them to [this case]". Another method is due to J.H. Cadwell (1947). For more details, see D S Mitrinović and J D Kečkić, *The Cauchy Method of Residues: Theory and Applications*, Dordrecht, etc: Reidel 1984, §5.3.4.10, pp. 158–168.

8. Recently T P Jameson (1994) (at the age of 16!) has suggested yet another method (subsequently suggested independently by S P Eveson (2005)). Consider the volume under the surface $z = \mathrm{e}^{-(x^2+y^2)}$, which is clearly given by

$$\begin{aligned} V &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{e}^{-(x^2+y^2)}\mathrm{d}x\mathrm{d}y \\ &= \left(\int_{-\infty}^{\infty} \mathrm{e}^{-x^2}\mathrm{d}x\right)^2. \end{aligned}$$

This can, however, also be thought of as a volume of revolution about the z axis where as $z = \mathrm{e}^{-x^2}$ we have $x = \sqrt{-\log z}$. Using the standard formula for a volume of revolution

$$V = \pi\int_0^1 x^2\mathrm{d}z = \pi\int_0^1 \{-\log z\}\,\mathrm{d}z = \pi[-z\log z + z]_0^1 = \pi.$$

and hence

$$\int_{-\infty}^{\infty} \mathrm{e}^{-x^2}\mathrm{d}x = \sqrt{\pi}.$$

References

(1) R C Archibald, A rare pamphlet of Moivre and some or his discoveries, *Isis* **8** (1926), 671–683.
(2) J C Burkill and H Burkill, *A Second Course in Mathematical Analysis*, Cambridge: University Press 1970.
(3) R Courant, *Differential and Integral Calculus* (2 volumes), London and Glasgow: Blackie 1934–6.
(4) Copson, *Theory of Functions of a Complex Variable*, Oxford: Clarendon Press 1935.
(5) A De Moivre, *Approximatio ad Summam Terminorum Binomii* $\overline{a+b}\backslash^n$ *in Seriem expansi* 1733, reprinted in Archibald (1929); translated with some additions in De Moivre (1738), De Moivre (1756) and Smith (1929).
(6) A De Moivre, *The Doctrine of Chances* (2nd ed.), London: H Woodfall 1738, reprinted London: Cass 1967.
(7) A De Moivre, *The Doctrine of Chances* (3rd ed.), London: A Millar 1756, reprinted New York, NY: Chelsea 1967 with a biographical article from *Scripta Mathematica* **2**(4) (1934), 316–333 by H M Walker.
(8) D Desbrow, Note 74.28 Evaluating the probability integral, *Mathematical Gazette* **74** (1990), 169–170,
(9) L Euler, *Calcul intégral*, translation of *Institvtionvm calcvli integralis*, Petropoli: Impensis Academiae Imperialis Scientiarum, 1768–1770.
(10) S P Eveson, Private communication (2005).
(11) T P Jameson, *Mathematical Gazette* **78** (1994), note 78.16, pp. 339–340.
(12) N Gauthier, Note 72.22 Evaluating the probability integral, *Mathematical Gazette*, **72** (1988), 124–125.
(13) B V Gnedenko, *Theory of Probability*, Moscow: Nauka 1954 and 1961, English translation Moscow: Mir 1969 and New York, NY: Chelsea 1967.
(14) A Hald, *A History of the Theory of Probability and Statistics and Their Application before 1750*, New York, NY, etc: Wiley 1990.
(15) P S Laplace, Mémoire sur la probabilité des causes par les évenments, *Mémoires de matheématique et de physique presentés à l'Académie royale des sciences, par divers savans, & lus dans ses assemblées* **6**, 621–626, reprinted in Laplace's *Oeuvres complètes* **8**, 27–65, translated with an introduction by S M Stigler, *Statistical Science* **1** (1986), 359–378.
(16) P S Laplace, *Théorie Analytiques des Probabilités*, Paris: Courcier 1812, reprinted Bruxelles: Culture et Civilisation 1967.
(17) D S Mitrinović and J D Kečkić, *The Cauchy Method of Residues: Theory and Applications*, Dordrecht, etc: D Reidel 1984.
(18) D E Smith, *A Source Book in Mathematics* (2 volumes), New York, NY: McGraw-Hill 1929, reprinted New York, NY: Dover 1959.
(19) J K F Sturm, *Cours d'Analyse de l'école polytechnique*, Paris: Mallet-Bachelier, 1857.
(20) E C Titchmarsh, *Theory of Functions*, Oxford: University Press 1932 and 1939.
(21) I Todhunter, *A History of the Mathematical Theory of Probability from the time of Pascal to that of Laplace*, Cambridge and London: Macmillan 1865, reprinted New York, NY: Chelsea 1949.
(22) G Valiron, *Cours d'Analyse Mathématique* (2 volumes), Paris: Masson 1947.
(23) E T Whittaker and G N Watson, *A Course of Modern Analysis*, Cambridge: University Press 1902, 1915, 1920 and 1927.

P.M.L.