INVERSION FORMULAS IN NORMAL VARIABLE MAPPING

By John Riordan

Bell Telephone Laboratories, New York

1. Summary. The two inversion formulas considered here arise from study of G. A. Campbell's work on the Poisson summation, which is described more fully in the introduction and in the main consists of finding a function or mapping of a variable connected with the summation in terms of a normal (Gaussian) variable g. More generally, this last is a process often called "normalization of the variable" and associated with the names of E. A. Cornish and R. A. Fisher. The mapping is two-way and the main inversion formula determines co-efficients for one way from those for the other, both sets of coefficients being descriptive of their mappings. More precisely if x is a given variable, g a Gaussian variable, y a parameter of the mapping, and the two mappings are

$$x = g + \sum_1^\infty G_n(g)\, y^n/n!,$$

$$g = x + \sum_1^\infty X_n(x) y^n/n!,$$

the formula expresses $G_n(x)$ in terms of $X_i(x)$, $i \leq n$, and vice versa.

The second formula is more particularly related to the Poisson summation and relates coefficients $p_n \equiv p_n(g)$ and $q_n \equiv q_n(g)$ in the pair of equations

$$a = c \sum_0^\infty q_n c^{-\frac{1}{2}n}/n!$$

$$c = a \sum_0^\infty p_n a^{-\frac{1}{2}n}/n!$$

Both formulas, which are necessarily elaborate, are given concise expression by the use of the multi-variable polynomials of E. T. Bell.

2. Introduction. In 1923, in a paper little known in statistical circles, G. A. Campbell [2] gave as the basis for his extensive tabulation of the Poisson summation an asymptotic series expressing the average a in terms of a normal variable g, corresponding to the probability of at least c occurrences, and c itself. That is to say, he associated with the Poisson summation

$$P(a, c) = \sum_c^\infty e^{-a} a^x/x!$$

a normal variable g, defined by

$$P(a, c) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{g} e^{-x^2/2}\, dx$$

and inverted the summation (which, as is well known, is equivalent to the incomplete Gamma function ratio) to give a series for a in terms of g and c. The series, which is carried to 11 terms, starts as follows:

$$a \sim c\left[1 + gc^{-1/2} + \frac{g^2 - 1}{3} c^{-1} + \frac{g^3 - 7g}{36} c^{-3/2} + \cdots\right]$$

If $x = (a - c)\, c^{-1/2}$ is introduced, this becomes

$$x \sim g + \frac{g^2 - 1}{3} c^{-1/2} + \frac{g^3 - 7g}{36} c^{-1} + \cdots$$

and x is seen to be, like g, a standardized variable of mean 0, variance 1.

It seems to have gone unnoticed that this result includes the χ^2 distribution through the transformation: $2a = X^2$, $2c = n$ and it has been rediscovered by A. M. Peiser [7] (4 terms) and by Goldberg and Levine [4] (6 terms).

It is possible also to express c in terms of a and g, and a formula of this kind with fewer terms which appears in a footnote in Campbell's paper is as follows:

$$c \sim a\left[1 - ga^{-1/2} + \frac{g^2 + 2}{6} a^{-1} + \frac{g^3 + 2g}{72} a^{-3/2} + \cdots\right]$$

Finally there is a third possibility of expressing g in terms of the remaining variables, preferably x and c; though unnoticed by Campbell this has since been brought to prominence by Cornish and Fisher [3], Hotelling and Frankel [5] and Kendall [6].

The idea behind the first expansion appears most clearly in the second form and is that for c large the variable x behaves nearly like g. The third possibility reverses this expansion and gives a function of x and c which behaves like g; hence if this function is first evaluated, reference to the normal integral table gives an immediate evaluation of the probabilities in question. Put in another way, the expansion widens the scope of the normal integral table and for this reason has been called "normalization" of the variable (but this term seems preempted by its use in another sense for orthogonal functions, and has been replaced in the title by normal variable mapping).

From the point of view of statistical theory, the three expressions are different versions of one relationship, which suggests that there should be general rules for transforming a series of one type into that of another. The two inversion formulas given below supply these rules in what appears to be as compact a form as the problem allows. It will be noted that the proofs given suppose convergent series, a case which leads to clarity and brevity and is interesting in itself. Applied to Campbell's series, they give the known results so far as the latter go, but of course for other asymptotic series they need independent verifications.

3. First Inversion Formula. This relates coefficients in series like Campbell's first and its reverse as in Cornish and Fisher. More precisely

If $G_1(g)$, $G_2(g)$ $\cdots$ *are assigned polynomials and if*

$$x = g + \sum_{n=1}^{\infty} G_n(g) y^n / n!, \tag{1}$$

defines x *in terms of* g *and a parameter* y, *then*

$$g = x + \sum_{1}^{\infty} X_n(x) y^n / n!, \tag{2}$$

where

$$-X_n(x) = Y_n(aG_1(x), aG_2(x), \cdots, aG_n(x)), \tag{3}$$

TABLE 1

Bell Polynomials Y_n $(fg_1, fg_2 \cdots fg_n)$

$$\begin{aligned}
Y_1 &= f_1 g_1 \\
Y_2 &= f_1 g_2 + f_2 g_1^2 \\
Y_3 &= f_1 g_3 + f_2(3g_2 g_1) + f_3 g_1^3 \\
Y_4 &= f_1 g_4 + f_2(4g_3 g_1 + 3g_2^2) + f_3(6g_2 g_1^2) + f_4 g_1^4 \\
Y_5 &= f_1 g_5 + f_2(5g_4 g_1 + 10 g_3 g_2) + f_3\,(10 g_3 g_1^2 + 15 g_2^2 g_1) \\
&\quad + f_4\,(10 g_2 g_1^3) + f_5 g_1^5 \\
Y_6 &= f_1 g_6 + f_2(6g_5 g_1 + 15 g_4 g_2 + 10 g_3^2) \\
&\quad + f_3(15 g_4 g_1^2 + 60 g_3 g_2 g_1 + 15 g_2^3) \\
&\quad + f_4(20 g_3 g_1^3 + 45 g_2^2 g_1^2) + f_5(15 g_2 g_1^4) + f_6 g_1^6 \\
Y_7 &= f_1 g_7 + f_2(7g_6 g_1 + 21 g_5 g_2 + 35 g_4 g_3) \\
&\quad + f_3(21 g_5 g_1^2 + 105 g_4 g_2 g_1 + 70 g_3^2 g_1 + 105 g_3 g_2^2) \\
&\quad + f_4(35 g_4 g_1^3 + 210 g_3 g_2 g_1^2 + 105 g_2^3 g_1) \\
&\quad + f_5(35 g_3 g_1^4 + 105 g_2^2 g_1^3) + f_6(21 g_2 g_1^5) + f_7 g_1^7 \\
Y_8 &= f_1 g_8 + f_2(8 g_7 g_1 + 28 g_6 g_2 + 56 g_5 g_3 + 35 g_4^2) \\
&\quad + f_3(28 g_6 g_1^2 + 168 g_5 g_2 g_1 + 280 g_4 g_3 g_1 + 210 g_4 g_2^2 + 280 g_3^2 g_2) \\
&\quad + f_4(56 g_5 g_1^3 + 420 g_4 g_2 g_1^2 + 280 g_3^2 g_1^2 + 840 g_3 g_2^2 g_1 + 105 g_2^4) \\
&\quad + f_5(70 g_4 g_1^4 + 560 g_3 g_2 g_1^3 + 420 g_2^3 g_1^2) \\
&\quad + f_6(56 g_3 g_1^5 + 210 g_2^2 g_1^4) + f_7(28 g_2 g_1^6) + f_8 g_1^8
\end{aligned}$$

Y_n *being the multivariable polynomial of* E. T. Bell [1], *in the variables* $G_1(x)$ *to* $G_n(x)$ *and the symbolic variable* a *which is such that*

$$a^i \equiv a_i = (-D)^{i-1}, \qquad D = d/dx,$$

with differentiations on all products of $G_1(x)$ *to* $G_n(x)$ *associated with it in the polynomial.*

Note the symmetry of x and g, which allows the transformation to go either way, the inverse of (3) being

$$-G_n(g) = Y_n(aX_1(g), aX_2(g) \cdots, aX_n(g)) \tag{4}$$

Table I gives explicit expressions for polynomials Y_1 to Y_8. It will be noted that the number of terms in Y_n is the number of partitions of n and that f_i, the variable replacing a_i in the table, is associated with terms corresponding to partitions with i parts; that is to say, if $Y_{n,i}$ designates such terms

$$Y_n = \sum_1^n f_i Y_{n,i}$$

The verification or extension of the table may be accomplished by the formulas and relations given by Bell (l.c.) or more directly by those modifications of Bell given by myself in [8].

The first few instances of (3), dropping the common variable x for brevity, may be read off from Table I (with appropriate changes of notation and interpretation of a_i) as follows:

$$-X_1 = G_1$$
$$-X_2 = G_2 - D(G_1^2)$$
$$-X_3 = G_3 - 3D(G_2G_1) + D^2(G_1^3)$$
$$-X_4 = G_4 - 4D(G_3G_1) - 3D(G_2^2) + 6D^2(G_2G_1^2) - D^3(G_1^4)$$

Applied to Campbell's first formula in its second form with $y = c^{-1/2}$ and

$$G_1(x) = (x^2 - 1)/3, \qquad G_3(x) = (-6x^4 - 14x^2 + 32)/270,$$
$$G_2(x) = (x^3 - 7x)/18, \qquad G_4(x) = (9x^5 + 256x^3 - 433x)/1680,$$

these show e.g.

$$-X_2 = \frac{x^3 - 7x}{18} - \frac{2(x^2 - 1)}{3} \cdot \frac{2x}{3} = \frac{-7x^3 + x}{18},$$

and similarly for the others, resulting in

$$X_1 = -(x^2 - 1)/3$$
$$X_2 = (7x^3 - x)/18$$
$$X_3 = -(219x^4 - 14x^2 - 13)/270$$
$$X_4 = (3993x^5 - 152x^3 + 119x)/1680$$

These determine a calculation formula for the Poisson summation, which is a refinement of the normal approximation. That is to say

$$P(a, c) = \Phi(g) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{g} e^{-t^2/2}\, dt$$

with

$$g = x - \frac{x^2 - 1}{3\sqrt{c}} + \frac{7x^3 - x}{36c} - \frac{219x^4 - 14x^2 - 13}{1620c\sqrt{c}} + \frac{3993x^5 - 152x^3 + 119x}{40320c^2} - \cdots$$

and $x = (a - c)/\sqrt{c}$.

For the t-variate, the formula is applied in the reverse direction since Hotelling and Frankel supply the first four values of X_n, that is, in present notation, the series

$$g \sim x - \frac{x^3 + x}{4} y + \frac{13x^5 + 8x^3 + 3x}{48} \frac{y^2}{2} - \frac{35x^7 + 19x^5 + x^3 - 15x}{64} \frac{y^3}{6} + \frac{6271x^9 + 3224x^7 - 102x^5 - 1680x^3 - 945x}{3840} \frac{y^4}{24} + \cdots$$

The reversed series (obtained by (4)) is

$$x \sim g + \frac{g^3 + g}{4} y + \frac{5g^5 + 16g^3 + 3g}{48} \frac{y^2}{2} + \frac{3g^7 + 19g^5 + 17g^3 - 15g}{64} \frac{y^3}{6} + \frac{79g^9 + 776g^7 + 1482g^5 - 1920g^3 - 945g}{3840} \frac{y^4}{24} + \cdots$$

The first three terms are checked by Goldberg and Levine (l.c.).

Another application worth noting is to the formulas of Cornish and Fisher which give $G_i(g)$ and $X_i(x)$ in terms of the relative cumulants of the distribution; to save space these are omitted.

The derivation of the formula may be indicated most easily by Lagrange's formula for the expansion of one function in powers of another in the following form[1]:

Let C be a contour in the complex z plane enclosing the point $z = x$, and let $f(z)$ and $\phi(z)$ be analytic on and inside C. Let y be such that $| y\phi(z) | < | z - x |$ when z is on C, and g be that root of the equation:

$$g = x + y\phi(g) \tag{5}$$

which lies inside C. Then

$$f(g) = \frac{1}{2\pi i} \int_C f(z) \frac{d}{dz} \{\log [z - x - y\phi(z)]\} \, dz = f(x) + \sum_1^\infty X_n^*(x) y^n/n! \tag{6}$$

where

$$X_n^*(x) = \frac{d^{n-1}}{dx^{n-1}} [f'(x)(\phi(x))^n] \tag{7}$$

The contour integral in (6) appears, slightly disguised, as a problem in Whittaker and Watson [*Modern Analysis*, Cambridge, 1920, p. 149]. The evaluation (7) is given for completeness, though no use is made of it in this section, the derivation proceeding directly from (6).

First notice that by (1) and (5)

$$-y\phi(g) = \sum_1^\infty G_n(g) y^n/n!,$$

[1] The author owes the suggestion for this to S. O. Rice, who also simplified the derivation of the second inversion formula given later.

so that the logarithm in (6) may be written

$$\log\,(z - x + \sum_1^\infty G_n(z)y^n/n!),$$

or

$$\log\,(z - x) + \log\left[1 + \sum_1^\infty G_n(z)(z - x)^{-1}y^n/n!\right],$$

or

$$\log\,(z - x) + \log \exp by, \tag{8}$$

with b a symbolic variable such that

$$b^0 \equiv b_0 = 1$$
$$b^n \equiv b_n = G_n(z)(z - x)^{-1}.$$

Now if

$$\begin{aligned} \log\,(\exp by) &= B_1y + B_2y^2/2! + \cdots, \\ &= \exp By, \end{aligned} \tag{9}$$

B being another symbolic variable, $B_0 = 0$, $B^n \equiv B_n$, it follows from equation (5) of [8] that

$$\begin{aligned} B_n &= [D_y^n \log\,(\exp by)]_{y=0}\,, \qquad D_y = d/dy, \\ &= Y_n(\beta b_1\,, \beta b_2\,, \cdots \beta b_n) \\ &= \sum_1^n \beta_i Y_{n,i}(b_1\,, b_2\,, \cdots b_n), \end{aligned}$$

with $\beta_i = (-)^{i-1}(i - 1)!$ and $Y_{n,i}$ the part of polynomial Y_n having i parts, as defined above. Moreover, each factor b_k of terms in $Y_{n,i}$ contributes $G_k(z)(z - x)^{-1}$ so that

$$B_n = \sum_1^n \beta_i(z - x)^{-i}\,Y_{n,i}(G_1(z), G_2(z) \cdots G_n(z)) \tag{11}$$

Then, by (5)

$$\begin{aligned} f(g) &= \frac{1}{2\pi i}\int_C f(z)\left(\frac{1}{z - x} + \frac{d}{dz}\exp By\right) dz \\ &= f(x) - \frac{1}{2\pi i}\int_C f'(z) \exp By\; dz \\ &= f(x) - \frac{1}{2\pi i}\int_C f'(z) \sum_1^\infty \frac{y^n}{n!} \sum_1^n \frac{\beta_i Y_{n,i}(G_1 \cdots G_n)}{(z - x)^i}\, dz \end{aligned}$$

$$= f(x) - \sum_1^\infty \frac{y^n}{n!} \sum_{i=1}^n \int_C \frac{(-)^{i-1}(i-1)!}{2\pi i(z-x)^i} Y_{n,i}(G_1(z) \cdots G_n(z))f'(z)\, dz$$

$$= f(x) - \sum_1^\infty \frac{y^n}{n!} \sum_1^n (-D)^{i-1}[f'(x)Y_{n,i}(G_1(x) \cdots G_n(x))]$$

with $D = d/dx$. The evaluation in the last line is by the Cauchy formula for derivates; the second line is derived by an integration by parts.

Equation (4) follows from this and the substitution $f(g) = g$.

4. Second Inversion Formula. This gives the interrelations of coefficients of series like the two Campbell series mentioned in the introduction. It runs as follows:

If $q_1(g)$, $q_2(g)$ $\cdots$ are given polynomials and if

$$a = c \sum_0^\infty \frac{q_n(g)c^{-\frac{1}{2}n}}{n!} \tag{12}$$

defines a in terms of g and a parameter c; then

$$c = a \sum_0^\infty \frac{p_n(g)a^{-\frac{1}{2}n}}{n!} \tag{13}$$

where

$$-p_n(g) = Y_n(\alpha q_1(g), \alpha q_2(g), \cdots, \alpha q_n(g)) \tag{14}$$

with $\alpha^1 \equiv \alpha_1 = 1;\ \alpha^i \equiv \alpha_i = (n-4)(n-6) \cdots (n-2i)2^{-i+1}$

Equation (14) is formally similar to (3) and by symmetry as before, $q_n(g)$ is readily expressible as a Y_n polynomial in $p_1(g)$ to $p_n(g)$.

The first five instances of (14), dropping the argument for brevity, are

$$-p_1 = q_1$$

$$-p_2 = q_2 - q_1^2$$

$$-p_3 = q_3 - \tfrac{3}{2} q_2q_1 + \tfrac{3}{4} q_1^3$$

$$-p_4 = q_4$$

$$-p_5 = q_5 + \tfrac{5}{2}(q_4q_1 + 2g_3q_2) - \tfrac{5}{4}(2q_3q_1^2 + 3q_2^2q_1) + \tfrac{15}{4} q_2q_1^3 - \tfrac{15}{16} q_1^5$$

Applied to Campbell's first series where

$$q_1(g) = g \qquad q_3(g) = (g^3 - 7g)/6$$

$$q_2(g) = \tfrac{2}{3}(g^2 - 1) \qquad q_4(g) = (-12g^4 - 28g^2 + 64)/135$$

$$q_5(g) = (36g^5 + 1024g^3 - 1732g)/1296$$

these show that

$$p_1(g) = -g \qquad p_3(g) = (g^3 + 2g)/12$$
$$p_2(g) = (g^2 + 2)/3 \qquad p_4(g) = (12g^4 + 28g^2 - 64)/135$$
$$p_5(g) = (207g^5 + 2596g^3 - 6148g)/1296$$

The proof of (14) is as follows. First, for brevity introduce symbolic variables p and q with the usual interpretation $p^n \equiv p_n(g)$, $q^n \equiv q_n(g)$ so that (12) and (13) read

$$a = c \exp q\, c^{-\frac{1}{2}}$$
$$c = a \exp p\, a^{-\frac{1}{2}}$$

Now write $a = 1/x^2$, $c = 1/y^2$ changing these to

$$x = y\,(\exp qy)^{-\frac{1}{2}}$$
$$y = x\,(\exp px)^{-\frac{1}{2}}$$

and note that

$$x^2 y^{-2} = (\exp qy)^{-1} = \exp px \tag{15}$$

which shows that p_n is the coefficient of $x^n/n!$ in the expansion in powers of x of $(\exp qy)^{-1}$. Lagrange's formula gives at once ($D = d/dy$):

$$f(y) = \sum_1^\infty \frac{x^n}{n!} D^{n-1}[f'(g)(\exp qy)^{\frac{1}{2}n}]_{y=0} \tag{16}$$

so that

$$\begin{aligned}(\exp qy)^{-1} &= \sum_1^\infty \frac{x^n}{n!} D^{n-1}[-(\exp qy)^{\frac{1}{2}(n-4)} D(\exp qy)]_{y=0}\\ &= \sum_1^\infty \frac{x^n}{n!} D^{n-1}\left[-\frac{2}{n-2} D(\exp qy)^{\frac{1}{2}(n-2)}\right]_{y=0}\\ &= \sum_1^\infty \frac{x^n}{n!}\left(\frac{-2}{n-2}\right)[D^n(\exp qy)^{\frac{1}{2}(n-2)}]_{y=0}\end{aligned}$$

or

$$\begin{aligned}-p_n &= \frac{2}{n-2}[D^n(\exp qy)^{\frac{1}{2}(n-2)}]_{y=0}\\ &= Y_n(\alpha q_1, \alpha q_2, \cdots, \alpha q_n)\end{aligned} \tag{17}$$

with α_i as in (14), by equation (5) of [8].

BIBLIOGRAPHY

[1] Bell, E. T., "Exponential polynomials," *Annals of Math.* Vol. *35*, (1934), pp. 258–277.
[2] Campbell, G. A., "Probability curves showing Poisson's exponential summation,"

Bell System Tech Jl. Vol. *2*, (1923), pp. 95–113; Collected Papers, New York, 1937, pp. 224–242.
[3] Cornish, E. A. and Fisher, R. A., "Moments and cumulants in the specification of distributions," *Revue de l'Institut Intern. de Stat.* Vol. *4*, (1937), pp. 1–14.
[4] Goldberg, H. and Levine, H., "Approximate formulas for the percentage points and normalization of t and χ^2," *Annals of Math. Stat.* Vol. *17*, (1946), pp. 216–225.
[5] Hotelling, H. and Frankel, L. R., "The transformation of statistics to simplify their distribution," *Annals of Math. Stat.* Vol. *9*, (1938), pp. 87–96.
[6] Kendall, M. G., *The Advanced Theory of Statistics* I, London. 1943.
[7] Peiser, A. M., "Asymptotic formulas for significance levels of certain distributions," *Annals of Math. Stat.* Vol. *14*, (1943), pp. 56–62.
[8] Riordan, John, "Derivatives of Composite Functions," *Bull. Am. Math. Soc.* Vol. *52*, (1946), pp. 664–667.