Differentiation by integration using orthogonal polynomials, a survey

Tom Koornwinder

University of Amsterdam, T.H.Koornwinder@uva.nl,

`http://www.science.uva.nl/~thk/`

Lecture on June 7, 2011 at the Hanoi Institute of Mathematics, Hanoi, Vietnam

Work in collaboration with Enno Diekema (Apeldoorn) and last part with Michael Schlosser (Vienna)

Last modified: June 7, 2011

Approximations of the derivative $f'(x)$ $\quad (\delta \downarrow 0)$

1. $\dfrac{f(x+\delta)-f(x)}{\delta} \;\to\; f'(x)$ and $\dfrac{f(x+\delta)-f(x-\delta)}{2\delta} \;\to\; f'(x)$.
2. Take $g(y) = ay + b$ such that
   $\sum_{j=-N}^{N} |f\,(x+\delta j) - g\,(x+\delta j)|^2$ is minimal. Then

$$a = \frac{3}{2N(N+\frac{1}{2})(N+1)\delta} \sum_{j=-N}^{N} f\,(x+\delta j)\; j \;\to\; f'(x)$$

3. Take $g(y) = ay + b$ such that $\int_{-1}^{1} \left|f(x+\delta t) - g(x+\delta t)\right|^2 dt$ is minimal. Then

$$a = \frac{3}{2\delta} \int_{-1}^{1} f(x+\delta t) t \, dt \;\to\; f'(x).$$

Note: For $N = 1$ item 2 reduces to item 1b.

Note: For δ in item 2 replaced bt δ/N we can see that a in item 2 tends to a in item 3 as $N \to \infty$.

Item 3 as limit of item 2 for $N \to \infty$

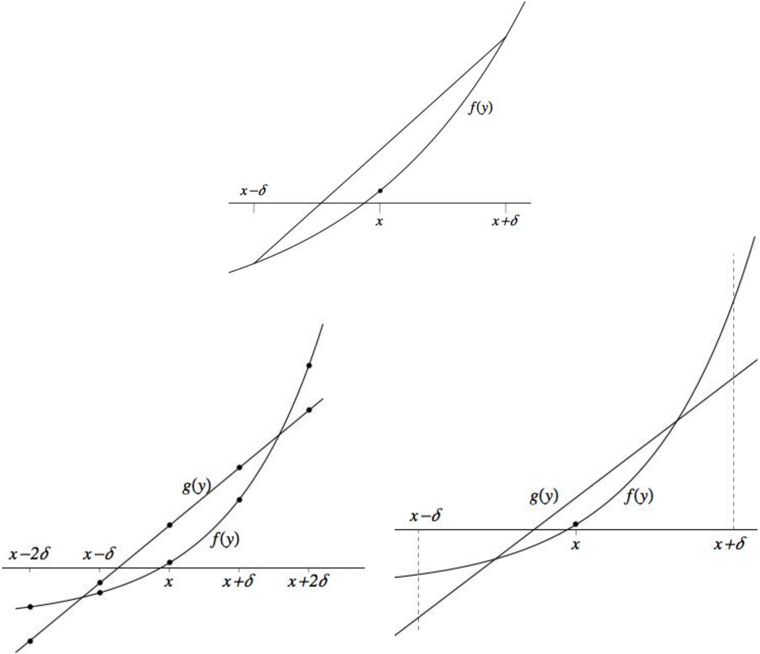
For $N \to \infty$ we have

$$\frac{1}{N}\sum_{j=-N}^{N}\left|f\left(x+\frac{\delta j}{N}\right)-g\left(x+\frac{\delta j}{N}\right)\right|^2 \to \int_{-1}^{1}\left|f(x+\delta t)-g(x+\delta t)\right|^2 dt$$

and

$$\frac{3}{2\delta}\,\frac{N^2}{(N+\frac{1}{2})(N+1)}\,\frac{1}{N}\sum_{j=-N}^{N} f\left(x+\frac{\delta j}{N}\right)\frac{j}{N} \to \frac{3}{2\delta}\int_{-1}^{1} f(x+\delta t)t\,dt$$

Graphs of f and g in the three cases



Transfer function

Suppose $D_\delta f$ is an approximation of $f'(x)$. The *transfer function* associated with D_δ is

$$H_\delta(\omega) := e^{-i\omega x} D_\delta(y \mapsto e^{i\omega y}).$$

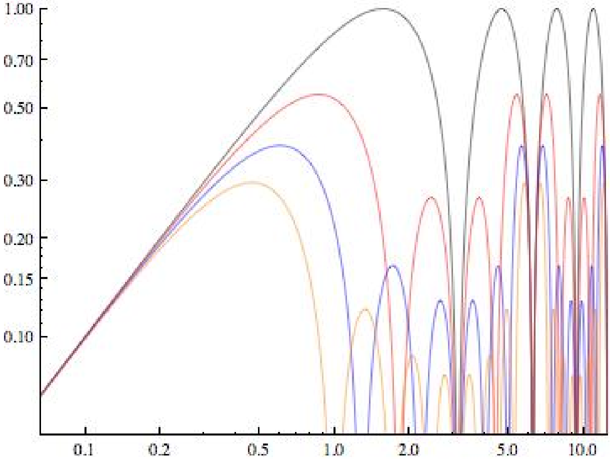
In the examples $H_\delta(\omega) = i\omega\, h(\delta\omega)$. Preferably, $|h|$ is close to 1 for small argument and close to 0 for larger argument (*low pass filter*).

For the previous three cases the transfer function can be computed to be:

1. $H_\delta(\omega) = i\omega\, e^{\frac{1}{2} i\omega\delta} \dfrac{\sin(\frac{1}{2}\omega\delta)}{\frac{1}{2}\omega\delta}$ and $H_\delta(\omega) = i\omega \dfrac{\sin(\omega\delta)}{\omega\delta}$.
2. $$H_\delta(\omega) = i\omega \frac{3}{2(2N+1)\sin^2(\delta\omega/2)\,\delta\omega} \times \left(\frac{\sin(N\delta\omega)}{N} - \frac{\sin((N+1)\delta\omega)}{N+1}\right).$$
3. $H_\delta(\omega) = i\omega \dfrac{3\big(\sin(\omega\delta) - \omega\delta\cos(\omega\delta)\big)}{(\omega\delta)^3}$.

log-log plots of absolute values of transfer functions

For $\dfrac{3}{2N(N+\frac{1}{2})(N+1)} \displaystyle\sum_{j=-N}^{N} f(x+j)\, j \qquad (N = 1, 2, 3, 4$ in black, red, blue and orange, repectively.



Orthogonal polynomials

$$\int_{\mathbb{R}} p_m(x)\, p_n(x)\, d\mu(x) = h_n\, \delta_{m,n}, \qquad p_n(x) = k_n\, x^n + \cdots$$

Let f be given. Minimize $\int_{\mathbb{R}} \left| f(x+\delta t) - g(t) \right|^2 d\mu(t)$ where g is a polynomial of degree $\le n$. Then g is the projection in $L^2(\mu)$ of f on the polynomials of degree $\le n$:

$$g(s) = \sum_{j=0}^{n} \frac{p_j(s)}{h_j} \int_{\mathbb{R}} f(x+\delta t)\, p_j(t)\, d\mu(t).$$

Then $\delta^{-n}\, g^{(n)}(0) \to f^{(n)}(x)$ as $\delta \downarrow 0$. Indeed,

$$\begin{aligned}
\delta^{-n}\, g^{(n)}(0) &= \frac{p_n^{(n)}(0)}{h_n\, \delta^n} \int_{\mathbb{R}} f(x+\delta t)\, p_n(t)\, d\mu(t) \\
&= \frac{k_n\, n!}{h_n\, \delta^n} \int_{\mathbb{R}} f(x+\delta t)\, p_n(t)\, d\mu(t).
\end{aligned}$$

Now plug in the Taylor series for $f(x+\delta t)$:

Taylor series

$$f(x+\delta t) = \sum_{j=0}^{n} f^{(j)}(x)\,\frac{(\delta t)^j}{j!} + o\big((\delta t)^n\big)$$

Then

$$\begin{aligned}
\delta^{-n}\, g^{(n)}(0) &= \frac{k_n\, n!}{h_n\, \delta^n} \int_{\mathbb{R}} f(x+\delta t)\, p_n(t)\, d\mu(t) \\
&= \frac{k_n\, n!}{h_n\, \delta^n} \sum_{j=0}^{n} \frac{\delta^j\, f^{(j)}(x)}{j!} \int_{\mathbb{R}} p_n(t)\, t^j\, d\mu(t) + o(1) \\
&= f^{(n)}(x)\, \frac{k_n}{h_n} \int_{\mathbb{R}} p_n(t)\, t^n\, d\mu(t) + o(1) \\
&= f^{(n)}(x) + o(1).
\end{aligned}$$

At least, if μ has bounded support. Otherwise assume that f has at most polynomial growth.

Similarly with Peano derivatives

Suppose that for certain c_k

$$f(y) = \sum_{k=0}^{n} \frac{c_k}{k!} (y - x)^k + o(|y - x|)^n) \quad \text{as } y \to x.$$

Then c_n is called the n-th *Peano derivative at x*. This was introduced by Peano in 1891.

Then the same approximation result holds, with $f^{(n)}(x)$ replaced by c_n.

Typical cases of orthogonal polynomials

- μ absolutely continuous: $d\mu(x) = w(x)\,dx$.
  For instance *Jacobi polynomials* $p_n(x) = P_n^{(\alpha,\beta)}(x)$ with $\alpha, \beta > -1$ and

$$\int_{-1}^{1} P_n^{(\alpha,\beta)}(x)\, P_m^{(\alpha,\beta)}(x)\, (1-x)^{\alpha}(1-x)^{\beta}\, dx = 0 \quad (n \neq m).$$

  In particular *Legendre polynomials* $P_n(x) = P_n^{(0,0)}(x)$ (constant weight function on $[-1, 1]$).
- μ discrete:

$$\sum_{x \in X} p_n(x)\, p_m(x)\, w(x) = 0 \quad (n \neq m)$$

  with X a discrete subset of $\mathbb{R}$.
  In particular, X may be finite: $X = \{x_0, x_1, \ldots, x_N\}$. Then only p_n for $n = 0, 1, \ldots, N$.

Special cases

- $\int_{\mathbb{R}} g(t)\,d\mu(t) = \int_{-1}^{1} g(t)\,dt$.

  Then $p_n(t) := P_n(t)$ (Legendre polynomial).

$$f^{(n)}(x) = \frac{(2n+1)!}{2^{n+1}n!} \lim_{\delta\downarrow 0} \frac{1}{\delta^n} \int_{-1}^{1} f(x+\delta t)\,P_n(t)\,dt.$$

- $\int_{\mathbb{R}} g(t)\,d\mu(t) = N^{-1}\sum_{j=-N}^{N} g(j/N)$.

  Then $p_n(t) = Q_n(-N\,t + N; 0, 0, 2N)$

  (special symmetric renormalized Hahn polynomials).

$$f^{(n)}(x) = \frac{(2n+1)!}{2^{n+1}\,n!}\,\frac{(2N)^{n+1}}{(2N+1)_{n+1}} \lim_{\delta\downarrow 0} \frac{1}{\delta^n}$$
$$\times \frac{1}{N} \sum_{j=-N}^{N} f(x+\delta j/N)\,Q_n\big(N(1-j/N); 0, 0, 2N\big).$$

For $n = 1$ we have $p_1(t) = t$ in both cases.

Connection with the continuous wavelet transform

The *continuous wavelet transform* Φ_g is defined by

$$(\Phi_g f)(a,b) := |a|^{-1/2} \int_{\mathbb{R}} f(t)\, \overline{g\big(a^{-1}(t-b)\big)}\, dt.$$

Here $a, b \in \mathbb{R}$, $a \neq 0$, andt the *wavelet* g is a nonzero function in $(L^1 \cap L^2)(\mathbb{R})$ such that $\int_{\mathbb{R}} g(t)\, dt = 0$.

Now, fix $n > 0$, let $d\mu(t) = w(t)\, dt$ with $t \mapsto t^k w(t)$ in $(L^1 \cap L^2)(\mathbb{R})$ for $k \leq n$. Then we have a wavelet

$$g(t) := p_n(t)\, w(t).$$

Then

$$\int_{\mathbb{R}} f(x + t\delta)\, p_n(t)\, w(t)\, dt = \delta^{-1/2}\, (\Phi_g f)(\delta, x).$$

and

$$\frac{k_n\, n!}{h_n\, \delta^{n-\frac{1}{2}}}\, (\Phi_g f)(\delta, x) \;\to\; f^{(n)}(x).$$

History

1. Cioranescu (1938) for $d\mu(x) = w(x)\,dx$. Not picked up .
2. Haslam-Jones (1953): Suppose the n-th Peano derivative c_n of f at x exists. Let J be a finite interval and let ν be a signed measure on J such that $\int_J t^k\,d\nu(t) = 0$ for $k < n$ and $\neq 0$ for $k = n$. (For instance $d\nu(t) = p_n(t)\,d\mu(t)$.) Then

$$c_n = \lim_{\delta \downarrow 0} \frac{n!}{\delta^n} \frac{\int_J f(x+\delta t)\,d\nu(t)}{\int_J t^n\,d\nu(t)}\,. \qquad \text{Not picked up.}$$

3. Lanczos (1956):

$$\frac{3}{2\delta} \frac{N^2}{(N+\frac{1}{2})(N+1)} \frac{1}{N} \sum_{j=-N}^{N} f\left(x + \frac{\delta j}{N}\right) \frac{j}{N} \;\to\; f'(x)$$

and for $N \to \infty$: $\quad \dfrac{3}{2\delta} \displaystyle\int_{-1}^{1} f(x+\delta t)t\,dt \;\to\; f'(x).$

A lot of follow-up: *Lanczos derivative*.

4. Savitzky & Golay (1964)

For low values of n they find the polynomial g of degree $\leq n$ such that $\sum_{j=-N}^{N} \left| f(x + \delta j/N) - g(j/N) \right|^2$ is minimal. Then, for $m \leq n$, $\delta^{-m} g^{(m)}(0)$ approximates $f^{m)}(x)$.

The authors were interested in spectroscopy. Their publication in the journal *Analytical Chemistry* had thousands of citations. The connection with orthogonal polynomials was later given by Meer & Weiss (1992).

This work suggests a multi-term extension of our general approximation theorem of the n-th derivative.

A multi-term variant

Let g still be the orthogonal projection in $L^2(\mu)$ of $t \mapsto f(x+\delta t)$ on the space of polynomials of degree $\le n$. So

$$g(s) = \sum_{j=0}^{n} \frac{p_j(s)}{h_j} \int_{\mathbb{R}} f(x+\delta t)\, p_j(t)\, d\mu(t).$$

But now let $m \le n$ and consider $\delta^{-m} g^{(m)}(0)$ as a possible approximation of $f^{(m)}(x)$. Indeed one can show that

$$\begin{aligned} \delta^{-m} g^{(m)}(0) &= \frac{1}{\delta^m} \sum_{j=m}^{n} \frac{1}{h_j} \left(\int_{\mathbb{R}} f(x+\delta t)\, p_j(t)\, d\mu(t) \right) p_j^{(m)}(0) \\ &= f^{(m)}(x) + o(\delta^{n-m}). \end{aligned}$$

For fixed m the approximation becomes better as n gets bigger. Write $D_\delta^m f := \delta^{-m} g^{(m)}(0)$.

Transfer function for the multi-term variant

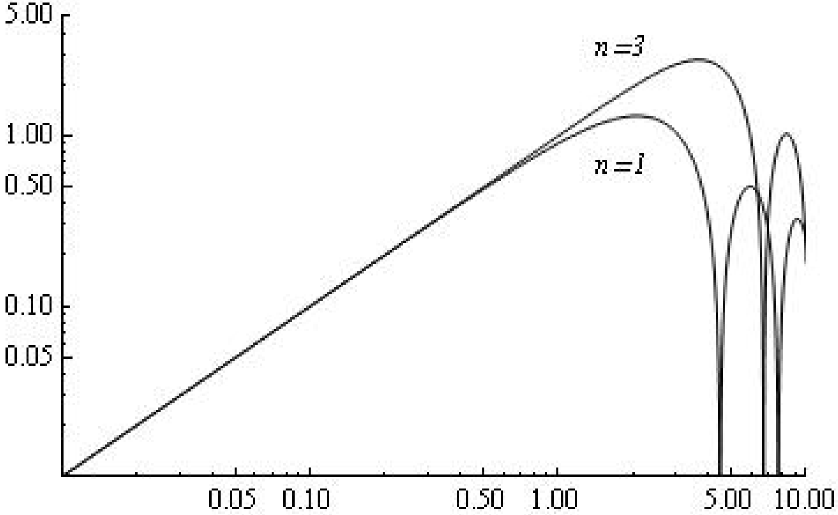
Moreover assume that μ is an even measure and that $n - m$ is odd. The transfer function is $H_\delta(\omega) := e^{-i\omega x}\, D_\delta^m(y \mapsto e^{i\omega y})$. Then, for some bounded function G:

$$H_\delta(\omega) = (i\omega)^m \Big(1 - \frac{|p_{n+1}^{(m)}(0)|}{|k_{n+1}|\,(n+1)!}\,(\delta\omega)^{n-m+1} + (\delta\omega)^{n-m+3}\, G(\delta\omega)\Big).$$

$|H_\delta(\omega)|$ stays close to ω^m for a while and then falls off. For fixed m the falling off starts at higher ω as n gets bigger.

So bigger n means better approximation but more perturbation by noise. A best suitable n has to be found depending on the particular problem.

Transfer function in Legendre case for $m = 1$ and $n = 1, 3$ (log-log plot)



The multi-term variant for $m = 0$: smoothing

$f(x)$ is aproximated by

$$\sum_{j=0}^{n} \frac{p_j(0)}{h_j} \int_{\mathbb{R}} f(x + \delta t)\, p_j(t)\, d\mu(t) = \int_{\mathbb{R}} f(x + \delta t)\, \mathbf{K}_n(t, 0)\, d\mu(t),$$

where

$$\mathbf{K}_n(t, s) = \sum_{j=0}^{n} \frac{p_j(t)\, p_j(s)}{h_j} = \frac{k_n}{k_{n+1}\, h_n} \, \frac{p_{n+1}(t) p_n(s) - p_n(t) p_{n+1}(s)}{t - s}$$

is the *Christoffel-Darboux kernel*.

Smoothing (earlier called *graduation*) has old nineteenth century roots, most in the finite discrete case. It was for instance important in actuarial sciences. We will change notation.

Smoothing in the finite discrete case

Let ρ be a function defined on $\{-N, -N+1, \ldots, N\}$ such that $\rho(-x) = \rho(x)$. We go from an input function f to an output function g by convolution with ρ:

$$g(y) = \sum_{x=-N}^{N} f(y-x)\,\rho(x).$$

In particular, take for some $n < N$

$$\rho(x) = \mathbf{K}_{2n}(x, 0)\, w(x) = \sum_{j=0}^{n} \frac{p_{2j} x\, p_{2j}(0)}{h_{2j}}\, w(x),$$

where the p_k are orthogonal polynomials on $\{-N, -N+1, \ldots, N\}$ with respect to even weights $w(x)$.

Then the smoothing reproduces the polynomials of degree $\leq 2n+1$, but not all of degree $2n+2$. Such smoothing is called *exact* for degree $2n+1$.

The characteristic function

For smoothing by convolution with ρ the *characteristic function* is given by

$$\phi(\omega) := \sum_{x=-N}^{N} \rho(x)\, e^{-ix\omega}.$$

The smoothing is exact for degree $2n+1$ iff $\phi(\omega) = 1 - a\omega^{2n+2} + \cdots$ for some $a \neq 0$.

Define ρ_m by

$$\sum_{x=-mN}^{mN} f(y-x)\, \rho_m(x) := \sum_{x_1,\ldots,x_m=-N}^{N} f(y-x_1-\ldots-x_m)\rho(x_1)\ldots\rho(x_m).$$

De Forest (1878): When can asymptotic behaviour of $\rho_m(x)$ for $m \to \infty$ be described?

Schoenberg (1948): Iff $|\phi(\omega)| < 1$ for $0 < \omega < 2\pi$.
Then the smoothing is called *stable*. Then a above is positive.

Greville's work

Greville (1966) established stability for $\rho(x)$ corresponding to the weights of certain special symmetrized Hahn polynomials:

$$w_\alpha(x) = \binom{\alpha + N + x}{N + x} \binom{\alpha + N - x}{N - x}$$

$$(x = -N, -N+1, \ldots, N-1, N)$$

for $\alpha = 0, 1, 2, \ldots$.

He also took the limit for $\alpha \to \infty$. Then

$$w(x) = \binom{2N}{N + x} \quad (x = -N, -N+1, \ldots, N-1, N),$$

the weights of special symmetrized Krawtchouk polynomials.

However, he did not explicitly use orthogonal polynomials.

The symmetric Krawtchouk case

On the one hand

$$\begin{aligned}\phi(\omega) &= \sum_{x=-N}^{N} e^{-ix\omega}\,\mathbf{K}_{2n}(x,0)\,w(x) = 1 - a\,\omega^{2n+2} + \cdots \\ &= 1 - (\sin^2\omega/2)^{n+1}\,P(\cos^2\omega/2)\end{aligned}$$

for some polynomial P of degree $N-n-1$.

On the other hand

$$\phi^{(k)}(\pi) = \sum_{x=-N}^{N} (-ix)^k\,(-1)^x\,\mathbf{K}_{2n}(x,0)\,w(x) = 0 \quad \text{if } k < 2N-2n$$

(use that $p_{2N}(x) = (-1)^{N+x}$). Hence

$$\phi(\omega) = (\cos^2\omega/2)^{N-n}\,Q(\sin^2\omega/2)$$

for some polynomial Q of degree n.

The symmetric Krawtchouk case (continued)

Thus P of degree $N-n-1$ and Q of degree n are related by

$$1 = z^{n+1} P(1-z) + (1-z)^{N-n} Q(z).$$

$Q(z) =$ power series of $(1-z)^{-N+n}$ in z truncated after term of z^n

$$= \sum_{k=0}^{n} \frac{(N-n)_k}{k!} z^k, \text{ so } Q(\sin^2 \omega/2) > 0 \text{ so } \phi(\omega) > 0 \quad (0 \le \omega < \pi).$$

$P(1-z)$

$=$ power series of z^{-n-1} in $1-z$ truncated after term of $(1-z)^{N-n-1}$

$$= \sum_{k=0}^{N-n-1} \frac{(n+1)_k}{k!} (1-z)^k \, .$$

Hence

$$1 = z^{n+1} \sum_{k=0}^{N-n-1} \frac{(n+1)_k}{k!} (1-z)^k + (1-z)^{N-n} \sum_{k=0}^{n} \frac{(N-n)_k}{k!} z^k.$$

(*identity with two incomplete binomial series*)

The identity with two incomplete binomial series

$$1 = z^n \sum_{k=0}^{m-1} \binom{n+k-1}{k} (1-z)^k + (1-z)^m \sum_{k=0}^{n-1} \binom{m+k-1}{k} z^k.$$

History, variations and extensions considered in Schlosser & K (2008), triggered by the occurrence of the identity in an unpublished manuscript (2007) by P. de Jong.

We then traced back the identity to Chaundy & Bullard (1960). Their proof:

$$\begin{aligned} 1 = z^n \sum_{k=0}^{m-1} & \binom{m+n-1}{k} z^{m-k-1}(1-z)^k \\ & + (1-z)^m \sum_{\ell=0}^{n-1} \binom{m+n-1}{\ell} z^\ell (1-z)^{n-\ell-1} \\ & \qquad = z^n P_{m,n}(1-z) + (1-z)^m P_{n,m}(z). \end{aligned}$$

Now reason as before (Q follows already by symmetry from P).

Earlier occurrences of the identity: Hering (1868)

A. G. Hering, *Summation der n ersten Glieder der binomischen Reihe mittelst der Theorie der hypergeometrischen Reihen*, Programm der Realschule in Chemnitz, 1868; JFM 01.0089.04.

Durch Substitution dieser Werthe in die obige Gleichung erhalten wir aus 3) des vorigen §:

$$(1-x)_n^{-m} = (1-x)^{-m} - (1-x)^{-m}\, x^n \left(1-\overline{1-x}\right)_m^{-n}. \qquad 1)$$

Earlier occurrences of the identity: de Moivre (1738)

A. de Moivre, *The doctrine of chances*, London, second ed., 1738.

But as there is a particular elegancy for the Sums of a finite number of Terms in thoſe Series whoſe Coefficients are figurate numbers beginning at Unity, I ſhall ſet down the *Canon* for thoſe Sums.

Let n denote the number of Terms whoſe Sum is to be found, and p the rank or order which thoſe figurate numbers obtain; then the Sum will be

$$\frac{1-x^n}{\overline{1-x}|^{p}} - \frac{nx^n}{\overline{1-x}|^{p-1}} - \frac{n.\overline{n+1}.x^n}{1.2.\overline{1-x}|^{p-2}}$$

$$- \frac{n.\overline{n+1}.\overline{n+2}.x^n}{1.2.3.\overline{1-x}|^{p-3}} - \frac{n.\overline{n+1}.\overline{n+2}.\overline{n+3}.x^n}{1.2.3.4.\overline{1-x}|^{p-4}}, \text{ \&c.}$$

which is to be continued till the number of Terms be $=p$.

The series with n terms of figurate numbers of order p is:

$$\sum_{k=0}^{n-1} \binom{p+k-1}{k} x^k.$$

Earlier occurrences of the identity: de Montmart (1713)

Pierre Renard de Montmort, *Essay d'analyse sur les jeux de hasard*, 2nd edition, 1713

AUTRE FORMULE.

191. L'ANALYSE m'a encore fourni une autre formule. Supposant les mêmes dénominations que ci-dessus, je trouve le sort de Pierre $= \frac{1 \times a^p b^0}{\overline{a+b}^p} + \frac{p \times a^p b^1}{\overline{a+b}^{p+1}} + \frac{p \cdot \overline{p+1} \times a^p b^2}{1.2.\overline{a+b}^{p+2}} + \frac{p.\overline{p+1}.\overline{p+2} \times a^p b^3}{1.2.3 \times \overline{a+b}^{p+3}} + \frac{p.\overline{p+1}.\overline{p+2}.\overline{p+3}.a^p b^4}{1.2.3.4\,\overline{a+b}^{p+4}} +$ &c. Et de même le sort de Paul $= \frac{1 \times b^q a^0}{\overline{a+b}^q} + \frac{q \times b^q a^1}{\overline{a+b}^{q+1}} + \frac{q.\overline{q+1}\, b^q a^2}{1.2 \times \overline{a+b}^{q+2}} + \frac{q.\overline{q+1}.\overline{q+2}\, b^q a^3}{1.2.3\,\overline{a+b}^{q+3}} + \frac{q.\overline{q+1}.\overline{q+2}.\overline{q+3}\, b^q a^4}{1.2.3.4\,\overline{a+b}^{q+4}} +$ &c.

idem, translated

Pierre Renard de Montmort, *Essay d'analyse sur les jeux de hasard*, 2nd edition, 1713, translated.

If we suppose that the number of chances that Pierre has to win each point, or if we wish that his strength be to that of Paul as a to b; we will have likewise the lot of Pierre by multiplying the terms of this series which are the coefficients of the power m, by the powers of a & of b which correspond to them (*art. 27*); thus the preceding series becomes $a^m b^0 + ma^{m-1}b + \frac{m.m-1}{1.2}a^{m-2}bb + \frac{m.m-1.m-2}{1.2.3}a^{m-3}b^3 + \frac{m.m-1.m-2.m-3}{1.2.3.4}a^{m-4}b^4 + \&c.$ which it is necessary to continue to the number of terms expressed by q, & to divide by $(a+b)^m$. The formula which designates the lot of Paul is $1 \times b^m a^0 + mb^{m-1}a + \frac{m.m-1}{1.2}b^{m-2}aa + \frac{m.m-1.m-2}{1.2.3}b^{m-3}a^3 + \frac{m.m-1.m-2.m-3}{1.2.3.4}b^{m-4}a^4 + \&c.$ continued to the number of terms expressed by p, & divided by $(a+b)^m$.

ANOTHER FORMULA.

191. The Analysis has again furnished me another formula. Supposing the same denominations as above, I find the lot of Pierre $= \frac{1 \times a^p b^0}{(a+b)^p} + \frac{p \times a^p b^1}{(a+b)^{p+1}} + \frac{p.p+1 \times a^p b^2}{1.2.(a+b)^{p+2}} + \frac{p.p+1.p+2 \times a^p b^3}{1.2.3.(a+b)^{p+3}} + \frac{p.p+1.p+2.p+3 \times a^p b^4}{1.2.3.4.(a+b)^{p+4}} + \&c.$ And likewise of Paul $= \frac{1 \times b^q a^0}{(a+b)^q} + \frac{q \times b^q a^1}{(a+b)^{q+1}} + \frac{q.q+1 \times b^q a^2}{1.2.(a+b)^{q+2}} + \frac{q.q+1.q+2 \times b^q a^3}{1.2.3.(a+b)^{q+3}} + \frac{q.q+1.q+2.q+3 \times b^q a^4}{1.2.3.4.(a+b)^{q+4}} + \&c.$

The formula which expresses the lot of Pierre will have as many terms as there are units in q, & that which expresses the lot of Paul as many terms as there are units in p.

The problem of points

This comes from a game of chance with two players Pierre and Paul who have chances p and $1-p$, respectively, of winning each round. The player who has first won a certain number of rounds will collect the entire prize. Suppose that the game is prematurely interrupted when Pierre has to win still n rounds and Paul m rounds. What is then a fair division of the stake?

First solution: Suppose $m+n-1$ rounds are still played.

Pierre: $\sum_{k=0}^{m-1} \binom{m+n-1}{k} p^{m+n-k-1}(1-p)^k$

Paul: $\sum_{\ell=0}^{n-1} \binom{m+n-1}{\ell} p^{\ell}(1-p)^{m+n-\ell-1}$

Second solution: Suppose they still play until there is a winner (at most $m+n-1$ rounds).

Pierre: $p \sum_{k=0}^{m-1} \binom{n+k-1}{k} p^{n-1}(1-p)^k$

Paul: $(1-p) \sum_{k=0}^{n-1} \binom{m+k-1}{k} p^k(1-p)^{m-1}$

Further reading

1. E. Diekema and T. H. Koornwinder, *Differentiation by integration using orthogonal polynomials, a survey*, `arXiv:1102.5219v1`.
2. T. H. Koornwinder and M.J. Schlosser, *On an identity by Chaundy and Bullard. I*, Indag. Math. (N.S.) 19 (2008), 239–261; `arXiv:0712.2125v3`.