Multiset

From Wikipedia, the free encyclopedia

In mathematics, a **multiset** (or **bag**) is a generalization of the concept of a set that, unlike a set, allows multiple instances of the multiset's elements. For example, $\{a, a, b\}$ and $\{a, b\}$ are different multisets although they are the same set. However, order does not matter, so $\{a, a, b\}$ and $\{a, b, a\}$ are the same multiset.

The **multiplicity** of an element is the number of instances of the element in a specific multiset. For example, an infinite number of multisets exist which contain only elements a and b, varying only by multiplicity:

- The unique set $\{a, b\}$ contains only elements a and b, each having multiplicity 1
- In multiset $\{a, a, b\}$, a has multiplicity 2 and b has multiplicity 1
- In multiset $\{a, a, a, b, b, b\}$, a and b both have multiplicity 3

Nicolaas Govert de Bruijn coined the word *multiset* in the 1970s, according to Donald Knuth.[1]:694 However, the use of multisets predates the word *multiset* by many centuries. Knuth attributes the first study of multisets to the Indian mathematician Bhāskarāchārya, who described permutations of multisets around 1150. Knuth also lists other names that were proposed or used for multisets, including *list*, *bunch*, *bag*, *heap*, *sample*, *weighted set*, *collection*, and *suite*.[1]:694

Contents

Overview

The number of times an element belongs to the multiset is the multiplicity of that member. The total number of elements in a multiset, including repeated memberships, is the cardinality of the multiset. For example, in the multiset $\{a, a, b, b, b, c\}$ the multiplicities of the members a, b, and c are respectively 2, 3, and 1, and the cardinality of the multiset is 6. To distinguish between sets and multisets, a notation that incorporates square brackets is sometimes used: the multiset $\{2, 2, 3\}$ can be represented as $[2, 2, 3]$.[2] In multisets, as in sets and in contrast to tuples, the order of elements is irrelevant: The multisets $\{a, a, b\}$ and $\{a, b, a\}$ are equal.

History

Wayne Blizard traced multisets back to the very origin of numbers, arguing that "in ancient times, the number n was often represented by a collection of n strokes, tally marks, or units."[3] These and similar collections of objects are multisets as strokes, tally marks, or units are considered indistinguishable. This shows that people implicitly used multisets even before mathematics emerged.

This shows that necessity in this structure has been always so urgent that multisets have been several times rediscovered and appeared in literature under different names.[4] For instance, they were referred to as bags by James Lyle Peterson in 1981,[5] who attributed this term to a 1971 paper by Cerf et al.[6] A multiset has been also called an aggregate, heap, bunch, sample, weighted set, occurrence set, and fireset (finitely repeated element set).[4][7]

Although multisets were implicitly utilized from ancient times, their explicit exploration happened much later. The first known study of multisets is attributed to the Indian mathematician Bhāskarāchārya circa 1150, who described permutations of multisets.[1]:694 The work of Marius Nizolius (1498–1576) contains another early reference to the concept of multisets.[8] Athanasius Kircher found the number of multiset permutations when one element can be repeated.[9] Jean Prestet published a general rule for multiset permutations in 1675.[10] John Wallis explained this rule in more detail in 1685.[11]

In the explicit form, multisets appeared in the work of Richard Dedekind.[12][13] Other mathematicians formalized multisets and began to study them as a precise mathematical object (structure) in the 20th century.

Example multisets

One of the simplest and most natural examples is the multiset of prime factors of a number n. Here the underlying set of elements is the set of prime divisors of n. For example, the number 120 has the prime factorization

$$120 = 2^3 3^1 5^1$$

which gives the multiset $\{2, 2, 2, 3, 5\}$.

A related example is the multiset of solutions of an algebraic equation. A quadratic equation, for example, has two solutions. However, in some cases they are both the same number. Thus the multiset of solutions of the equation could be $\{3, 5\}$, or it could be $\{4, 4\}$. In the latter case it has a solution of multiplicity 2.

A special case of the above are the eigenvalues of a matrix, if these are defined as the multiset of roots of its characteristic polynomial. However a choice is made here: the (usual) definition of eigenvalues does not refer to the characteristic polynomial, and other possibilities give rise to different notions of multiplicity, and therefore to different multisets. The geometric multiplicity of λ as eigenvalue of a matrix A is the dimension of the kernel of $A - \lambda I$, which is often smaller than its multiplicity as root of the characteristic polynomial (the algebraic multiplicity) when the latter is larger than 1. The set of eigenvalues of A is also the set of roots of its minimal polynomial, but the multiset of those roots need not be the same either as the one defined using algebraic multiplicity, or using the geometric multiplicity.

Formal definition

Within set theory, a **multiset** may be formally defined as a 2-tuple (A, m) where A is some set and $m : A \to \mathbb{N}_{\geq 1}$ is a function from A to the set $\mathbb{N}_{\geq 1} = \{1, 2, 3, \ldots\}$ of positive integers. The set A is called the *underlying set of elements*. For each a in A the *multiplicity* (that is, number of occurrences) of a is the number $m(a)$. If a universe U in which the elements of A must live is specified, the definition can be simplified to just a multiplicity function $m_U : U \to \mathbb{N}_{\geq 0}$ from U to the set $\mathbb{N}_{\geq 0} = \{0, 1, 2, 3, \ldots\}$ of nonnegative integers, obtained by extending m to U with values 0 for each element not in A. This extended multiplicity function is

the multiplicity function called 1_A below. Like any function, the function m may be defined as its graph: the set of ordered pairs $\{(a, m\,(a)) : a \in A\}$. With these definitions the multiset written as $\{a, a, b\}$ is defined as $(\{a, b\}, \{(a, 2), (b, 1)\})$, and the multiset $\{a, b\}$ is defined as $(\{a, b\}, \{(a, 1), (b, 1)\})$.

The concept of a multiset is a generalization of the concept of a set. A multiset corresponds to an ordinary set if the multiplicity of every element is one (as opposed to some larger natural number). An indexed family, (a_i), where i is in some index-set, may define a multiset, sometimes written $\{a_i\}$, in which the multiplicity of any element x is the number of indices i such that $a_i = x$. The condition for this to be possible is that no element occurs infinitely many times in the family: even in an infinite multiset, the multiplicities must be finite numbers.

It is possible to extend the definition of a multiset by allowing multiplicities of individual elements to be infinite cardinals instead of natural numbers. Not all properties carry over to this generalization however, and this article uses the definition above, with finite multiplicities.

Multiplicity function

The set indicator function of a normal set is generalized to the multiplicity function for multisets. The **set indicator function** of a subset A of a set X is the function

$$1_A : X \to \{0, 1\}$$

defined by

$$1_A(x) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \notin A. \end{cases}$$

The set indicator function of the intersection of sets is the minimum function of the indicator functions

$$1_{A \cap B}(x) = \min\{1_A(x), 1_B(x)\}.$$

The set indicator function of the union of sets is the maximum function of the indicator functions

$$1_{A \cup B}(x) = \max\{1_A(x), 1_B(x)\}.$$

The set indicator function of the set difference of sets is the zero-truncated subtraction of the indicator functions

$$1_{A \setminus B}(x) = \max(0, 1_A(x) - 1_B(x)).$$

The set indicator function of a subset is smaller than or equal to that of the superset

$$A \subseteq B \Leftrightarrow \forall x,\, 1_A(x) \le 1_B(x).$$

The set indicator function of a cartesian product is the product of the indicator functions of the cartesian factors

$$1_{A \times B}(x, y) = 1_A(x) \cdot 1_B(y).$$

The cardinality of a (finite) set is the sum of the indicator function values

$$|A| = \sum_{x \in X} 1_A(x).$$

Now generalize the concept of set indicator function by releasing the constraint that the values are 0 and 1 only and allow the values 2, 3, 4 and so on. The resulting function is called a **multiplicity function** and such a function defines a **multiset**. The concepts of intersection, union, difference, subset, cartesian product and

cardinality of multisets are defined by the above formulas. There are also the following additional operations.

The multiplicity function of a **multiset sum**, is the sum of the multiplicity functions

$$\mathbf{1}_{A \uplus B}(x) = \mathbf{1}_A(x) + \mathbf{1}_B(x).$$

The scalar multiplication of a multiset by a natural number n may be defined as:

$$\mathbf{1}_{n \otimes A}(x) = n \times \mathbf{1}_A(x).$$

Example multiset operations

These examples demonstrate these operations of multiplicity functions.

$$\{1,1,1,3\} \cap \{1,1,2\} = \{1,1\}$$
$$\{1,1\} \cup \{1,2\} = \{1,1,2\}$$
$$\{1,1\} \setminus \{1,2\} = \{1\}$$
$$\{1,1\} \subseteq \{1,1,1,2\}$$
$$\{1,1\} \times \{1,2\} = \{(1,1),(1,1),(1,2),(1,2)\}$$
$$|\{1,1\}| = 2$$
$$\{1,1\} \uplus \{1,2\} = \{1,1,1,2\}$$
$$3 \otimes \{1,1,2\} = \{1,1,1,1,1,1,2,2,2\}$$

Free commutative monoids

The free commutative monoid on a set X (see free object) can be taken to be the set of finite multisets with elements drawn from X, with the monoid operation being multiset sum and the empty multiset as identity element. Such monoids are also known as (finite) formal sums of elements of X with natural coefficients. The free commutative semigroup is the subset of the free commutative monoid which contains all multisets with elements drawn from X except the empty multiset.

Free abelian groups are formal sums (i.e. linear combinations) of elements of X with integer coefficients. Equivalently, they may be seen as *signed* finite multisets with elements drawn from X.

Counting multisets

The number of multisets of cardinality k, with elements taken from a finite set of cardinality n, is called the **multiset coefficient** or **multiset number**. This number is written by some authors as $\left(\!\binom{n}{k}\!\right)$, a notation that is meant to resemble that of binomial coefficients; it is used for instance in (Stanley, 1997), and could be pronounced "n multichoose k" to resemble "n choose k" for $\binom{n}{k}$. Unlike for binomial coefficients, there is no "multiset theorem" in which multiset coefficients would occur, and they should not be confused with the unrelated multinomial coefficients that occur in the multinomial theorem.

The value of multiset coefficients can be given explicitly as

$$\left(\!\!\binom{n}{k}\!\!\right) = \binom{n+k-1}{k} = \frac{(n+k-1)!}{k!\,(n-1)!} = \frac{n(n+1)(n+2)\cdots(n+k-1)}{k!},$$

where the second expression is as a binomial coefficient; many authors in fact avoid separate notation and just write binomial coefficients. So, the number of such multisets is the same as the number of subsets of cardinality k in a set of cardinality $n + k - 1$. The analogy with binomial coefficients can be stressed by writing the numerator in the above expression as a rising factorial power

$$\left(\!\!\binom{n}{k}\!\!\right) = \frac{n^{\overline{k}}}{k!},$$

to match the expression of binomial coefficients using a falling factorial power:

$$\binom{n}{k} = \frac{n^{\underline{k}}}{k!}.$$

There are for example 4 **multisets** of cardinality 3 with elements taken from the set $\{1, 2\}$ of cardinality 2 ($n = 2$, $k = 3$), namely $\{1, 1, 1\}$, $\{1, 1, 2\}$, $\{1, 2, 2\}$, $\{2, 2, 2\}$. There are also 4 **subsets** of cardinality 3 in the set $\{1, 2, 3, 4\}$ of cardinality 4 ($n + k - 1$), namely $\{1, 2, 3\}$, $\{1, 2, 4\}$, $\{1, 3, 4\}$, $\{2, 3, 4\}$.

One simple way to prove the equality of multiset coefficients and binomial coefficients given above, involves representing multisets in the following way. First, consider the notation for multisets that would represent
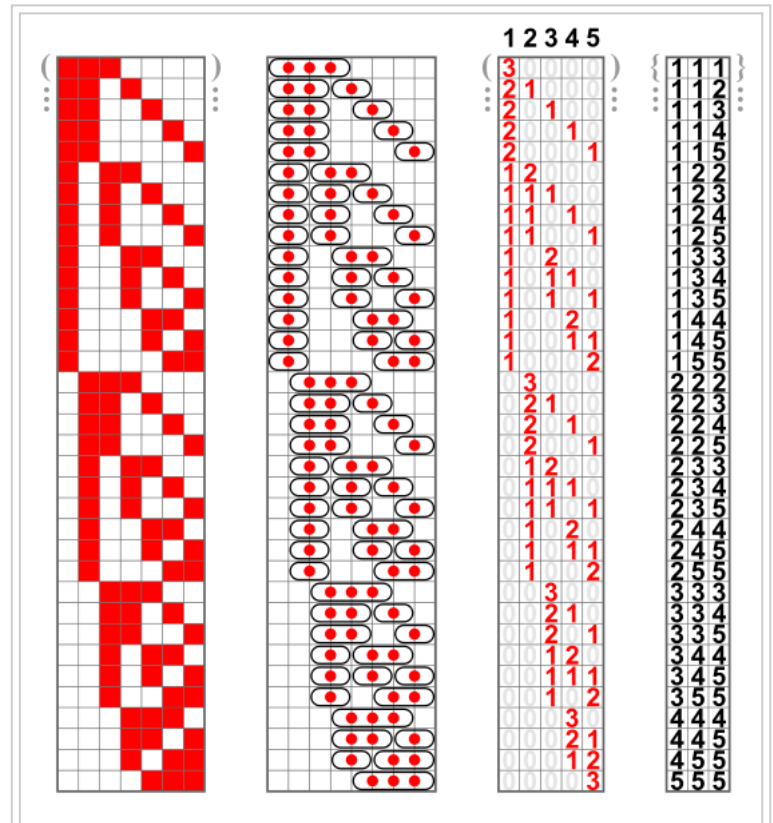
Bijection between 3-subsets of a 7-set (left) and 3-multisets with elements from a 5-set (right)
So this illustrates that $\binom{7}{3} = \left(\!\!\binom{5}{3}\!\!\right)$.

$\{a, a, a, a, a, a, b, b, c, c, c, d, d, d, d, d, d, d\}$ (6 as, 2 bs, 3 cs, 7 ds) in this form:

$$\bullet\bullet\bullet\bullet\bullet\bullet \mid \bullet\bullet \mid \bullet\bullet\bullet \mid \bullet\bullet\bullet\bullet\bullet\bullet\bullet$$

This is a multiset of cardinality $k = 18$ made of elements of a set of cardinality $n = 4$. The number of characters including both dots and vertical lines used in this notation is $18 + 4 - 1$. The number of vertical lines is $4 - 1$. The number of multisets of cardinality 18 is then the number of ways to arrange the $4 - 1$ vertical lines among the $18 + 4 - 1$ characters, and is thus the number of subsets of cardinality $4 - 1$ in a set of cardinality $18 + 4 - 1$. Equivalently, it is the number of ways to arrange the 18 dots among the $18 + 4 - 1$ characters, which is the number of subsets of cardinality 18 of a set of cardinality $18 + 4 - 1$. This is

$$\binom{4 + 18 - 1}{4 - 1} = \binom{4 + 18 - 1}{18} = 1330,$$

thus is the value of the multiset coefficient and its equivalencies:

$$\left(\!\!\binom{4}{18}\!\!\right) = \binom{21}{18} = \frac{21!}{18!\,3!} = \binom{21}{3} = \left(\!\!\binom{19}{3}\!\!\right),$$

$$= \frac{4 \cdot 5 \cdot 6 \cdot 7 \cdot 8 \cdot 9 \cdot 10 \cdot 11 \cdot 12 \cdot 13 \cdot 14 \cdot 15 \cdot 16 \cdot 17 \cdot 18 \cdot \mathbf{19 \cdot 20 \cdot 21}}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 7 \cdot 8 \cdot 9 \cdot 10 \cdot 11 \cdot 12 \cdot 13 \cdot 14 \cdot 15 \cdot 16 \cdot 17 \cdot 18},$$

$$= \frac{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdots 16 \cdot 17 \cdot 18 \cdot \mathbf{19 \cdot 20 \cdot 21}}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdots 16 \cdot 17 \cdot 18 \times \mathbf{1 \cdot 2 \cdot 3}},$$

$$= \frac{19 \cdot 20 \cdot 21}{1 \cdot 2 \cdot 3}.$$

One may define a generalized binomial coefficient

$$\binom{n}{k} = \frac{n(n-1)(n-2)\cdots(n-k+1)}{k!}$$

in which n is not required to be a nonnegative integer, but may be negative or a non-integer, or a non-real complex number. (If $k = 0$, then the value of this coefficient is 1 because it is the empty product.) Then the number of multisets of cardinality k in a set of cardinality n is

$$\left(\!\!\binom{n}{k}\!\!\right) = (-1)^k \binom{-n}{k}.$$

Recurrence relation

A recurrence relation for multiset coefficients may be given as

$$\left(\!\!\binom{n}{k}\!\!\right) = \left(\!\!\binom{n}{k-1}\!\!\right) + \left(\!\!\binom{n-1}{k}\!\!\right) \quad \text{for } n, k > 0$$

with

$$\left(\!\!\binom{n}{0}\!\!\right) = 1, \quad n \in \mathbb{N}, \quad \text{and} \quad \left(\!\!\binom{0}{k}\!\!\right) = 0, \quad k > 0.$$

The above recurrence may be interpreted as follows. Let $[n] := \{1, \cdots, n\}$ be the source set. There is always exactly one (empty) multiset of size 0, and if $n = 0$ there are no larger multisets, which gives the initial conditions.

Now, consider the case in which $n, k > 0$. A multiset of cardinality k with elements from $[n]$ might or might not contain any instance of the final element n. If it does appear, then by removing n once, one is left with a multiset of cardinality $k - 1$ of elements from $[n]$, and every such multiset can arise, which gives a total of

$$\left(\!\!\binom{n}{k-1}\!\!\right) \text{ possibilities.}$$

If n does not appear, then our original multiset is equal to a multiset of cardinality k with elements from $[n - 1]$, of which there are

$$\left(\!\!\binom{n-1}{k}\!\!\right).$$

Thus,

$$\left(\!\!\binom{n}{k}\!\!\right) = \left(\!\!\binom{n}{k-1}\!\!\right) + \left(\!\!\binom{n-1}{k}\!\!\right).$$

Generalization and connection to the negative binomial series

The multiplicative formula allows the definition of multiset coefficients to be extended by replacing n by an arbitrary number α (negative, real, complex):

$$\left(\!\!\binom{\alpha}{k}\!\!\right) = \frac{\alpha^{\overline{k}}}{k!} = \frac{\alpha(\alpha+1)(\alpha+2)\cdots(\alpha+k-1)}{k(k-1)(k-2)\cdots 1} \quad \text{for } k \in \mathbb{N} \text{ and arbitrary } \alpha.$$

With this definition one has a generalization of the negative binomial formula (with one of the variables set to 1), which justifies calling the $\left(\!\!\binom{\alpha}{k}\!\!\right)$ negative binomial coefficients:

$$(1-X)^{-\alpha} = \sum_{k=0}^{\infty} \left(\!\!\binom{\alpha}{k}\!\!\right) X^k.$$

This Taylor series formula is valid for all complex numbers α and X with $|X| < 1$. It can also be interpreted as an identity of formal power series in X, where it actually can serve as definition of arbitrary powers of series with constant coefficient equal to 1; the point is that with this definition all identities hold that one expects for exponentiation, notably

$$(1-X)^{-\alpha}(1-X)^{-\beta} = (1-X)^{-(\alpha+\beta)} \quad \text{and} \quad ((1-X)^{-\alpha})^{-\beta} = (1-X)^{-(-\alpha\beta)},$$

and formulas such as these can be used to prove identities for the multiset coefficients.

If α is a nonpositive integer n, then all terms with $k > -n$ are zero, and the infinite series becomes a finite sum. However, for other values of α, including positive integers and rational numbers, the series is infinite.

Polynomial notation

The set $\{x\}$ may be represented by the monomial x. Its power set, $\{\{\}, \{x\}\}$, is represented by the binomial $1 + x$.

The set $\{x, y\}$ may be represented by the monomial xy. Its power set, $\{\{\}, \{x\}, \{y\}, \{x, y\}\}$, is represented by the polynomial

$$(1 + x)(1 + y) = 1 + x + y + xy.$$

The multiset $\{x, x\}$ may be represented by the monomial x^2. The multiset of submultisets of $\{x, x\}$, $\{\{\}, \{x\}, \{x\}, \{x, x\}\}$, is represented by the polynomial

$$(1 + x)^2 = 1 + 2x + x^2.$$

The multiset of submultisets of the multiset represented by the polynomial x^n is

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

That is why the binomial coefficient counts the number of k-combinations of an n-set.

The multiset represented by the polynomial $x^K y^{N-K}$, containing N elements, K of which are hits, is called a statistical population, and a submultiset is called a statistical sample. The set of samples is

$$(1 + x)^K (1 + y)^{N-K}$$

which by the binomial theorem equals

$$\sum_{n=0}^{N} \sum_{k=0}^{K} \binom{K}{k} \binom{N-K}{n-k} x^k y^{n-k}.$$

So the number of n-samples with k hits is

$$\binom{K}{k}\binom{N-K}{n-k}.$$

See hypergeometric distribution and inferential statistics for further on the distribution of hits.

The infinite set of finite multisets of elements taken from the set $\{x\}$:

$$\{\{\},\{x\},\{x,x\},\{x,x,x\},\cdots\}$$

may be represented by the formal power series

$$S = 1 + x + x^2 + x^3 + \cdots = 1 + xS.$$

The formal solution,

$$S = (1-x)^{-1},$$

makes sense as a set of multisets, but the intermediate result, $1 - x$, does not make sense as a set of multisets.

The infinite set of finite multisets of elements taken from the set represented by the monomial xy is

$$(1-x)^{-1}(1-y)^{-1} = 1 + (x+y) + (x^2 + xy + y^2) + \cdots.$$

The infinite multiset of finite multisets of elements taken from the multiset represented by the monomial x^2 may be taken as a special case where $x = y$:

$$(1-x)^{-2} = 1 + 2x + 3x^2 + \cdots.$$

The infinite multiset of finite multisets of elements taken from the multiset represented by the monomial x^n is

$$(1-x)^{-n} = \sum_{k=0}^{\infty}\binom{-n}{k}(-x)^k.$$

This explains why "multisets are negative sets." The negative binomial coefficients count the number of k-multisets of elements taken from an n-set.

Cumulant generating function

A non-negative integer n, can be represented by the monomial x^n. A finite **multiset of non-negative integers** can likewise be represented by a polynomial $f(x)$.

It is convenient to consider the **cumulant generating function** $g(t) = \log f(e^t)$.

- The **cardinality** of the multiset is $e^{g(0)} = f(1)$.
- The derivative of the cumulant generating function is $g'(t) = f(e^t)^{-1}f'(e^t)e^t$.
    - The mean value of the multiset is $\mu = g'(0) = f(1)^{-1}f'(1)$.
    - The variance of the multiset is $\sigma^2 = g''(0)$.

For example, the multiset $\{2, 2, 2, 3, 5\}$ can be represented by the polynomial $f(x) = 3x^2 + x^3 + x^5$, with cumulant generating function $g(t) = \log(3e^{2t} + e^{3t} + e^{5t})$, cardinality $3 + 1 + 1 = 5$, derivative $g'(t) = (3e^{2t} + e^{3t} + e^{5t})^{-1}(6e^{2t} + 3e^{3t} + 5e^{5t})$, and mean value $\mu = (3 + 1 + 1)^{-1}(6 + 3 + 5) = 2.8$.

The numbers $(\mu, \sigma^2, \cdots) = (g'(0), g''(0), \cdots)$ are called cumulants.

The infinite set of non-negative integers $\{0, 1, 2, \cdots\}$ is represented by the formal power series $1 + x + x^2 + \cdots = (1 - x)^{-1}$. The mean value and standard deviation are undefined. Nevertheless, it has a cumulant-generating function, $g(t) = -\log(1 - e^t)$. The derivative of this cumulant-generating function is $g'(t) = (e^{-t} - 1)^{-1}$.

A finite **multiset of real numbers**, $A = \{A_i\}$, is represented by the cumulant generating function

$$g_A(t) = \log\left(\sum_i e^{tA_i}\right).$$

This representation is unique: different multisets have different cumulant generating functions. See partition function for the case where the numbers in question are the energy levels of a physical system.

The cumulant-generating function of a multiset of n real numbers having mean μ and standard deviation σ is $g(t) = \log n + \mu t + 2^{-1}(\sigma t)^2 + \cdots$ and the derivative is $g'(t) = \mu + \sigma^2 t + \cdots$.

The cumulant-generating function of a set $\{k\}$ consisting of a single real number k is $g(t) = kt$. The derivative is the number itself: $g'(t) = k$. So the concept of *the derivative of the cumulant generating function of a multiset of real numbers* is a generalization of the concept of a *real number*.

The cumulant-generating function of a constant multiset, $\{k, k, k, k, \cdots, k\}$ of n elements all equal to the same real number k, is $g(t) = \log n + kt$, and the derivative is the number itself: $g'(t) = k$, irrespective of n.

The cumulant-generating function of the multiset of sums of elements of two multisets of numbers is the sum of the two cumulant-generating functions:

$$
\begin{aligned}
g_{A+B}(t) &= \log\left(\sum_i \sum_j e^{t(A_i + B_j)}\right) \\
&= \log\left(\sum_i \sum_j e^{tA_i} e^{tB_j}\right) \\
&= \log\left(\sum_i e^{tA_i} \cdot \sum_j e^{tB_j}\right) \\
&= \log\left(\sum_i e^{tA_i}\right) + \log\left(\sum_j e^{tB_j}\right) \\
&= g_A(t) + g_B(t).
\end{aligned}
$$

There is no general formula for computing the cumulant generating function of a product

$$g_{A \cdot B}(t) = \log\left(\sum_i \sum_j e^{tA_i B_j}\right),$$

but the special case of a constant times a multiset of numbers is:

$$g_{k \cdot A}(t) = \log\left(\sum_i e^{t(kA_i)}\right)$$

$$= \log\left(\sum_i e^{(tk)A_i}\right)$$

$$= g_A(kt).$$

The multiset $2 \cdot A = \{2A_i\}$ is not the same multiset as $2 \times A = A + A = \{A_i + A_j\}$. For example, $2 \cdot \{1, -1\} = \{2, -2\}$ while $2 \times \{1, -1\} = \{1, -1\} + \{1, -1\} = \{1 + 1, 1 - 1, -1 + 1, -1 -1\} = \{2, 0, 0, -2\}$. The cumulant generating function of $k \times A$ is

$$g_{k \times A}(t) = kg_A(t).$$

The standard normal distribution is like a limit of large multisets of numbers.

$$\lim_{k \to \infty} k^{-1}\left(k^2 \times \{1, -1\}\right).$$

This limit does not make sense as a multiset of numbers, but the derivative of the cumulant generating functions of the multisets in question makes sense, and the limit is well defined.

$$\lim_{k \to \infty} g'_{k^{-1}(k^2 \times \{1,-1\})}(t) = \lim_{k \to \infty} \frac{d}{dt}\left(k^2 \log\left(e^{+tk^{-1}} + e^{-tk^{-1}}\right)\right)$$

$$= \lim_{k \to \infty} \frac{d}{dt}\left(k^2 \log 2 + 2^{-1}t^2 + \cdots\right)$$

$$= t.$$

The constant term $k^2 \log(2)$ vanishes by differentiation. The terms $\cdots$ vanish in the limit. So for the standard normal distribution, having mean 0 and standard deviation 1, the derivative of the cumulant generating function is $g'(t) = t$. For the normal distribution having mean μ and standard deviation σ, the derivative of the cumulant generating function is $g'(t) = \mu + \sigma^2 t$.

Applications

Multisets have various applications.[7] They are becoming the main structure of combinatorics because in its search for higher rigorousness, modern combinatorics has been developed not for sets but for multisets.[14][15][16][17] Multisets have become an important tool in databases.[18][19][20] For instance, multisets are often used to implement relations in database systems. Multisets also play an important role in computer science.[1]

There are also other applications. For instance, Richard Rado used multisets as a device to investigate the properties of families of sets. He wrote, "The notion of a set takes no account of multiple occurrence of any one of its members, and yet it is just this kind of information which is frequently of importance. We need only think of the set of roots of a polynomial f(x) or the spectrum of a linear operator."[4]

Generalizations

Different generalizations of multisets have been introduced, studied and applied to solving problems.

- Fuzzy multisets[21]
- Rough multisets[22]
- Real-valued multisets (in which multiplicity of an element can be any real number)[23][24]

- Hybrid sets[25]
- where multiplicity is any real-valued step function[26]
- Soft multisets[27]
- Soft fuzzy multisets[28]
- Named set (unification of all generalizations of sets)[29][30][31][32]

See also

- Set Theory
- 🏛 Partitions of multisets

References

1. Knuth, Donald E. (1998). *Seminumerical Algorithms*. The Art of Computer Programming **2** (3rd ed.). Addison Wesley. ISBN 0-201-89684-2.
2. Hein, James L. (2003). *Discrete mathematics*. Jones & Bartlett Publishers. pp. 29–30. ISBN 0-7637-2210-3.
3. Blizard, Wayne D (1989). "Multiset theory". *Notre Dame Journal of Formal Logic* **30** (1): 36–66. doi:10.1305/ndjfl/1093634995
4. Blizard, Wayne D. (1991). "The Development of Multiset Theory". *Modern Logic*. **1** (4): 319–352.
5. Peterson, James L. (1981). *Petri Net Theory and the Modelling of Systems*. Englewood Cliffs, New Jersey: Prentice-Hall.
6. Cerf, Vint; Fernandez, E.; Gostelow, K.; Volansky, S. (December 1971). Formal Control Flow Properties of a Model of Computation (Report). Los Angeles, California: Computer Science Department, University of California. ENG-7178.
7. Singh, D.; Ibrahim, A. M.; Yohanna, T.; Singh, J. N. (2007). "An overview of the applications of multisets". *Novi Sad Journal of Mathematics* **37** (2): 73–92.
8. Angelelli, I. (1965). "Leibniz's misunderstanding of Nizolius' notion of 'multitudo'". *Notre Dame Journal of Formal Logic* (6): 319–322.
9. Kircher, Athanasius (1650). *Musurgia Universalis*. Rome: Corbelletti.
10. Prestet, Jean (1675). *Elemens des Mathematiques*. Paris: André Pralard.
11. Wallis, John (1685). *A treatise of algebra*. London: John Playford.
12. Dedekind, Richard (1888). *Was sind und was sollen die Zahlen?* Braunschweig: Vieweg.
13. Syropoulos, Apostolos (2001). "Mathematics of Multisets". In Calude, C. S.; et al. *Multiset processing: Mathematical, computer science, and molecular computing points of view*. Springer-Verlag. pp. 347–358.
14. Aigner, M. (1979). *Combinatorial Theory*. New York/Berlin: Springer Verlag.
15. Anderson, I. (1987). *Combinatorics of Finite Sets*. Oxford: Clarendon Press.
16. Stanley, Richard P. (1997). *Enumerative Combinatorics* **1**. Cambridge University Press. ISBN 0-521-55309-1.
17. Stanley, Richard P. (1999). *Enumerative Combinatorics* **2**. Cambridge University Press. ISBN 0-521-56069-1.
18. Grumbach, S.; Milo, T (1996). "Towards tractable algebras for bags". *Journal of Computer and System Sciences* **52** (3): 570–588. doi:10.1006/jcss.1996.0042
19. Libkin, L.; Wong, L. (1994). "Some properties of query languages for bags". *Proceedings of the Workshop on Database Programming Languages*. Springer Verlag. pp. 97–114.
20. Libkin, L.; Wong, L. (1995). "On representation and querying incomplete information in databases with bags". *Information Processing Letters*. **56** (4): 209–214. doi:10.1016/0020-0190(95)00154-5
21. Yager, R. R. (1986). "On the Theory of Bags". *International Journal of General Systems* **13**: 23–37. doi:10.1080/03081078608934952
22. Grzymala-Busse, J. (1987). "Learning from examples based on rough multisets". *Proceedings of the 2nd International Symposium on Methodologies for Intelligent Systems*. Charlotte, North Carolina. pp. 325–332.
23. Blizard, Wayne D. (1989). "Real-valued Multisets and Fuzzy Sets". *Fuzzy Sets and Systems* **33**: 77–97. doi:10.1016/0165-0114(89)90218-2
24. Blizard, Wayne D. (1990). "Negative Membership". *Notre Dame Journal of Formal Logic* **31** (1): 346–368.
25. Loeb, D. (1992). "Sets with a negative numbers of elements". *Advances in Mathematics* **91**: 64–74. doi:10.1016/0001-8708(92)90011-9.
26. Miyamoto, S. (2001). "Fuzzy Multisets and their Generalizations". *Multiset Processing*. **2235**: 225–235.
27. Alkhazaleh, S.; Salleh, A. R.; Hassan, N. (2011). "Soft Multisets Theory". *Applied Mathematical Sciences* **5** (72): 3561–3573.
28. Alkhazaleh, S.; Salleh, A. R. (2012). "Fuzzy Soft Multiset Theory". *Abstract and Applied Analysis*.
29. Burgin, Mark (1990). "Theory of Named Sets as a Foundational Basis for Mathematics". *Structures in Mathematical Theories*. San Sebastian. pp. 417–420.

30. Burgin, Mark (1992). "On the concept of a multiset in cybernetics". *Cybernetics and System Analysis* **3**: 165–167.
31. Burgin, Mark (2004). "Unified Foundations of Mathematics". arXiv:math/0403186.
32. Burgin, Mark (2011). *Theory of Named Sets*. Mathematics Research Developments. Nova Science Pub Inc. ISBN 978-1-61122-788-8.

Retrieved from "https://en.wikipedia.org/w/index.php?title=Multiset&oldid=767885068"

Categories: Basic concepts in set theory │ Factorial and binomial topics