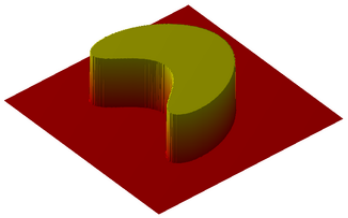
WIKIPEDIA

Indicator function

In mathematics, an **indicator function** or a **characteristic function** is a function defined on a set X that indicates membership of an element in a subset A of X, having the value 1 for all elements of A and the value 0 for all elements of X not in A. It is usually denoted by a symbol 1 or I, sometimes in boldface or blackboard boldface, with a subscript specifying the subset.



A three-dimensional plot of an indicator function, shown over a square two-dimensional domain (set X): the 'raised' portion overlays those two-dimensional points which are members of the 'indicated' subset (A).

In other contexts, such as computer science, this would more often be described as a **boolean predicate function** (to test set inclusion).

The Dirichlet function is an example of an indicator function and is the indicator of the rationals.

Contents

Definition

The indicator function of a subset A of a set X is a function

$$\mathbf{1}_A\colon X \to \{0,1\}$$

defined as

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

The Iverson bracket allows the equivalent notation, $[x \in A]$, to be used instead of $\mathbf{1}_A(x)$.

The function $\mathbf{1}_A$ is sometimes denoted I_A, χ_A, K_A or even just A. (The Greek letter χ appears because it is the initial letter of the Greek word χαρακτήρ, which is the ultimate origin of the word *characteristic*.)

The set of all indicator functions on X can be identified with $\mathcal{P}(X)$, the power set of X. Consequently, both sets are sometimes denoted by 2^X. This is a special case ($Y = \{0,1\} = 2$) of the notation Y^X for the set of all functions $f : X \to Y$.

Remark on notation and terminology

The notation χ_A is also used to denote the characteristic function in convex analysis, which is defined as if using the reciprocal of the standard definition of the indicator function.

A related concept in statistics is that of a dummy variable. (This must not be confused with "dummy variables" as that term is usually used in mathematics, also called a bound variable.)

The term "characteristic function" has an unrelated meaning in classic probability theory. For this reason, traditional probabilists use the term **indicator function** for the function defined here almost exclusively, while mathematicians in other fields are more likely to use the term *characteristic function* to describe the function that indicates membership in a set.

In fuzzy logic and modern many-valued logic, predicates are the characteristic functions of a probability distribution. That is, the strict true/false valuation of the predicate is replaced by a quantity interpreted as the degree of truth.

Basic properties

The *indicator* or *characteristic* function of a subset A of some set X, maps elements of X to the range {0,1}.

This mapping is surjective only when A is a non-empty proper subset of X. If $A \equiv X$, then $\mathbf{1}_A = 1$. By a similar argument, if $A \equiv \emptyset$ then $\mathbf{1}_A = 0$.

In the following, the dot represents multiplication, 1·1 = 1, 1·0 = 0 etc. "+" and "−" represent addition and subtraction. "∩" and "∪" is intersection and union, respectively.

If A and B are two subsets of X, then

$$\mathbf{1}_{A\cap B} = \min\{\mathbf{1}_A, \mathbf{1}_B\} = \mathbf{1}_A \cdot \mathbf{1}_B,$$
$$\mathbf{1}_{A\cup B} = \max\{\mathbf{1}_A, \mathbf{1}_B\} = \mathbf{1}_A + \mathbf{1}_B - \mathbf{1}_A \cdot \mathbf{1}_B,$$

and the indicator function of the complement of A i.e. A^C is:

$$\mathbf{1}_{A^\complement} = 1 - \mathbf{1}_A.$$

More generally, suppose $A_1, \ldots, A_n$ is a collection of subsets of X. For any $x \in X$:

$$\prod_{k\in I}(1 - \mathbf{1}_{A_k}(x))$$

is clearly a product of 0s and 1s. This product has the value 1 at precisely those $x \in X$ that belong to none of the sets A_k and is 0 otherwise. That is

$$\prod_{k\in I}(1 - \mathbf{1}_{A_k}) = \mathbf{1}_{X-\bigcup_k A_k} = 1 - \mathbf{1}_{\bigcup_k A_k}.$$

Expanding the product on the left hand side,

$$\mathbf{1}_{\bigcup_k A_k} = 1 - \sum_{F\subseteq\{1,2,\ldots,n\}} (-1)^{|F|}\mathbf{1}_{\bigcap_F A_k} = \sum_{\emptyset\neq F\subseteq\{1,2,\ldots,n\}} (-1)^{|F|+1}\mathbf{1}_{\bigcap_F A_k}$$

where $|F|$ is the cardinality of F. This is one form of the principle of inclusion-exclusion.

As suggested by the previous example, the indicator function is a useful notational device in combinatorics. The notation is used in other places as well, for instance in probability theory: if X is a probability space with probability measure $\mathbf{P}$ and A is a measurable set, then $\mathbf{1}_A$ becomes a random variable whose expected value is equal to the probability of A:

$$\mathrm{E}(\mathbf{1}_A) = \int_X \mathbf{1}_A(x)\,d\mathrm{P} = \int_A d\mathrm{P} = \mathrm{P}(A).$$

This identity is used in a simple proof of Markov's inequality.

In many cases, such as order theory, the inverse of the indicator function may be defined. This is commonly called the generalized Möbius function, as a generalization of the inverse of the indicator function in elementary number theory, the Möbius function. (See paragraph below about the use of the inverse in classical recursion theory.)

Mean, variance and covariance

Given a probability space $(\Omega, \mathcal{F}, \mathrm{P})$ with $A \in \mathcal{F}$, the indicator random variable $\mathbf{1}_A \colon \Omega \rightarrow \mathbb{R}$ is defined by $\mathbf{1}_A(\omega) = 1$ if $\omega \in A$, otherwise $\mathbf{1}_A(\omega) = 0$.

Mean

$$\mathrm{E}(\mathbf{1}_A(\omega)) = \mathrm{P}(A)$$

Variance

$$\mathrm{Var}(\mathbf{1}_A(\omega)) = \mathrm{P}(A)(1 - \mathrm{P}(A))$$

Covariance

$$\mathrm{Cov}(\mathbf{1}_A(\omega), \mathbf{1}_B(\omega)) = \mathrm{P}(A \cap B) - \mathrm{P}(A)\,\mathrm{P}(B)$$

Characteristic function in recursion theory, Gödel's and Kleene's *representing function*

Kurt Gödel described the *representing function* in his 1934 paper "On Undecidable Propositions of Formal Mathematical Systems". (The paper appears on pp. 41–74 in Martin Davis ed. *The Undecidable*):

> "There shall correspond to each class or relation R a representing function $\varphi(x_1, \ldots, x_n) = 0$ if $R(x_1, \ldots, x_n)$ and $\varphi(x_1, \ldots, x_n) = 1$ if $\sim R(x_1, \ldots, x_n)$." (p. 42; the "~" indicates logical inversion i.e. "NOT")

Stephen Kleene (1952) (p. 227) offers up the same definition in the context of the primitive recursive functions as a function φ of a predicate P takes on values 0 if the predicate is true and 1 if the predicate is false.

For example, because the product of characteristic functions $\varphi_1 * \varphi_2 * \ldots * \varphi_n = 0$ whenever any one of the functions equals 0, it plays the role of logical OR: IF $\varphi_1 = 0$ OR $\varphi_2 = 0$ OR . . . OR $\varphi_n = 0$ THEN their product is 0. What appears to the modern reader as the representing function's logical inversion, i.e. the representing function is 0 when the function R is "true" or satisfied", plays a useful role in Kleene's definition of the logical functions OR, AND, and IMPLY (p. 228), the bounded- (p. 228) and unbounded- (p. 279ff) mu operators (Kleene (1952)) and the CASE function (p. 229).

Characteristic function in fuzzy set theory

In classical mathematics, characteristic functions of sets only take values 1 (members) or 0 (non-members). In fuzzy set theory, characteristic functions are generalized to take value in the real unit interval [0, 1], or more generally, in some algebra or structure (usually required to be at least a poset or lattice). Such generalized characteristic functions are more usually called membership functions, and the corresponding "sets" are called *fuzzy* sets. Fuzzy sets model the gradual change in the membership degree seen in many real-world predicates like "tall", "warm", etc.

Derivatives of the indicator function

A particular indicator function is the Heaviside step function. The Heaviside step function $H(x)$ is the indicator function of the one-dimensional positive half-line, i.e. the domain $[0, \infty)$. The distributional derivative of the Heaviside step function is equal to the Dirac delta function, i.e.

$$\delta(x) = \tfrac{dH(x)}{dx},$$

with the following property:

$$\int_{-\infty}^{\infty} f(x)\,\delta(x)dx = f(0).$$

The derivative of the Heaviside step function can be seen as the 'inward normal derivative' at the 'boundary' of the domain given by the positive half-line. In higher dimensions, the derivative naturally generalises to the inward normal derivative, while the Heaviside step function naturally generalises to the indicator function of some domain D. The surface of D will be denoted by S. Proceeding, it can be derived that the inward normal derivative of the indicator gives rise to a 'surface delta function', which can be indicated by $\delta_S(\mathbf{x})$:

$$\delta_S(\mathbf{x}) = -\mathbf{n}_x \cdot \nabla_x \mathbf{1}_{\mathbf{x}\in D}$$

where n is the outward normal of the surface S. This 'surface delta function' has the following property:[1]

$$-\int_{\mathbf{R}^n} f(\mathbf{x})\,\mathbf{n}_x \cdot \nabla_x \mathbf{1}_{\mathbf{x}\in D}\; d^n\mathbf{x} = \oint_S f(\beta)\; d^{n-1}\beta.$$

By setting the function f equal to one, it follows that the inward normal derivative of the indicator integrates to the numerical value of the surface area S.

See also

- Dirac measure
- Laplacian of the indicator
- Dirac delta
- Extension (predicate logic)
- Free variables and bound variables
- Heaviside step function
- Iverson bracket
- Kronecker delta, a function that can be viewed as an indicator for the identity relation
- Macaulay brackets
- Multiset
- Membership function
- Simple function
- Dummy variable (statistics)
- Statistical classification
- Zero-one loss function

Notes

1. Lange, Rutger-Jan (2012), "Potential theory, path integrals and the Laplacian of the indicator", *Journal of High Energy Physics*, **2012** (11): 29–30, arXiv:1302.0864 (https://arxiv.org/abs/1302.0864), Bibcode:2012JHEP...11..032L (https://ui.adsabs.harvard.edu/abs/2012JHEP...11..032L), doi:10.1007/JHEP11(2012)032 (https://doi.org/10.1007%2FJHEP11%282012%29032)

References

- Folland, G.B. (1999). *Real Analysis: Modern Techniques and Their Applications* (Second ed.). John Wiley & Sons, Inc. ISBN 978-0-471-31716-6.
- Cormen, Thomas H.; Leiserson, Charles E.; Rivest, Ronald L.; Stein, Clifford (2001). "Section 5.2: Indicator random variables". *Introduction to Algorithms* (Second ed.). MIT Press and McGraw-Hill. pp. 94–99. ISBN 978-0-262-03293-3.
- Davis, Martin, ed. (1965). *The Undecidable*. New York: Raven Press Books, Ltd.
- Kleene, Stephen (1971) [1952]. *Introduction to Metamathematics* (Sixth Reprint with corrections). Netherlands: Wolters-Noordhoff Publishing and North Holland Publishing Company.
- Boolos, George; Burgess, John P.; Jeffrey, Richard C. (2002). *Computability and Logic*. Cambridge UK: Cambridge University Press. ISBN 978-0-521-00758-0.
- Zadeh, Lotfi A. (June 1965). "Fuzzy sets" (https://web.archive.org/web/20070622151801/http://www-bisc.cs.berkeley.edu/zadeh/papers/Fuzzy%20Sets-1965.pdf) (PDF). *Information and Control*. **8** (3): 338–353. doi:10.1016/S0019-9958(65)90241-X (https://doi.org/10.1016%2FS0019-9958%2865%2990241-X). Archived from the original (http://www-bisc.cs.berkeley.edu/zadeh/papers/Fuzzy%20Sets-1965.pdf) (PDF) on 2007-06-22.
- Goguen, Joseph (1967). "*L*-fuzzy sets". *Journal of Mathematical Analysis and Applications*. **18** (1): 145–174. doi:10.1016/0022-247X(67)90189-8 (https://doi.org/10.1016%2F0022-247X%2867%2990189-8).