On the concept of Bernoulliness

Vladimir Vovk

практические выводы
теории вероятностей
могут быть обоснованы
в качестве следствий
гипотез о *предельной*
при данных ограничениях
сложности изучаемых явлений

On-line Compression Modelling Project (New Series)

Working Paper #15

First posted December 28, 2016. Last revised June 23, 2024.

Project web site:
http://alrw.net

Abstract

The first part of this paper is another English translation of [17]. It gives a natural definition of a finite Bernoulli sequence (i.e., a typical realization of a finite sequence of binary IID trials) and compares it with the Kolmogorov–Martin-Löf definition, which is interpreted as defining exchangeable sequences. The appendix gives the historical background and proofs.

Contents

On the concept of Bernoulliness

This note gives a definition of a "Bernoulli sequence", i.e., a finite sequence of 0s and 1s that is random with respect to the class of Bernoulli measures. Our definition is different from A. N. Kolmogorov's [1] and more similar to the definition in [2]; for terminological convenience, sequences random in the sense of [1] will be called collectives.

1. Denote by X the union of the following sets: 2^*, the set of all finite sequences of 0s and 1s, N^*, the set of all finite sequences of natural numbers, N, the set of all natural numbers, and $\mathcal{F}(N)$, the set of all finite subsets of the set N (a finite subset of N is identified with the list of its elements in increasing order). The set X defined in this way contains all objects that we will need.

A *description method* is a partial function $B : 2^* \times X \to X$ that has an algorithm computing its values (the reader not familiar with the theory of algorithms can safely rely on his intuitive idea of computability). The length of a shortest p such that $B(p,y) = x$ will be denoted $K_B(x \mid y)$ and called the *complexity of x given y under the description method B.*

Lemma 1. *There exists a description method $A(p,y)$ such that, for any description method $B(p,y)$,*

$$K_A(x \mid y) \leq K_B(x \mid y) + C,$$

where C is a constant that does not depend on x and y.

A description method is *prefix* if, for any $p \in 2^*$ and $p' \in 2^*$ such that p is a prefix of p', $B(p,y) = B(p',y)$ for all $y \in X$. Lemma 1 will remain true if "description method" is replaced by "prefix description method".

Let us fix a description method A satisfying Lemma 1; $K_A(x \mid y)$ will be denoted $K(x \mid y)$ and called the *complexity of x given y*. The *prefix complexity* $KP(x \mid Y)$ *of x given y* is defined similarly. Proofs of the assertions made above can be found in [3].

2. Denote by $2^{(n)}$ the set of all sequences in 2^* of length n. Let $p \in [0,1]$ and $n > 0$ be an integer. On the set $2^{(n)}$ define the Bernoulli measure with parameters (n,p) as follows: for any $x \in 2^{(n)}$ set $P\{x\} = p^k(1-p)^{n-k}$, where k is the number of 1s in x. On the set $\{0,1,\ldots,n\}$ define the binomial measure with parameters (n,p) by the equality $P\{k\} = \binom{n}{k} p^k(1-p)^{n-k}$ for all $k = 0,1,\ldots,n$.

With each Bernoulli (binomial) measure P with parameters (n,p) associate an integer-valued function $T(x \mid P)$ of the variable $x \in 2^{(n)}$ (in the case of a binomial measure, $x \in \{0,1,\ldots,n\}$) so that:

(1) $E2^{T(x|P)} \leq 1$, where E stands for the mean under the measure P.

(2) As function of x and P, the function T is lower semicomputable. This means that there exists an algorithm A such that: given n, x, and an "oracle" that for each $i \in N$ outputs a rational number a_i satisfying

$|a_i - p| \le 2^{-i}$ (of course, the sequence a_i does not have to be computable), A enumerates a nondecreasing sequence of integer numbers m_j such that $\sup_j m_j = T(x \mid P)$; it is required that A should work correctly for an arbitrary "oracle" (for oracular computability, see [4]).

Such functions T will be called *tests* (for randomness). A condition similar to (1) first appeared in [5].

Lemma 2. *There exists a test $D(x \mid P)$ such that, for any test $T(x \mid P)$,*

$$D(x \mid P) \ge T(x \mid P) - C,$$

where C does not depend on x and P.

The test $D(x \mid P)$ in Lemma 2 will be called the *randomness deficiency of x with respect to the measure P*.

3. The *Bernoulliness deficiency* $\mathrm{D}^{\mathrm{Bernoul}}(x)$ of a sequence $x \in 2^{(n)}$ is defined to be $\inf D(x \mid P)$, where the inf is over the Bernoulli measures with parameters (n, p) for all $p \in [0, 1]$. Similarly, for a number k in $\{0, 1, \dots, n\}$ define the *binomiality deficiency* $\mathrm{D}^{\mathrm{binom}}_n(k)$ as $\inf D(k \mid P)$; here the inf is over the binomial measures. A sequence in 2^* is called *Bernoulli* if its Bernoulliness deficiency is small. We speak of a binomial number in a similar sense.

Theorem 1. *If $x \in 2^{(n)}$ contains k 1s, then*

$$\mathrm{D}^{\mathrm{Bernoul}}(x) - \left[\log_2 \binom{n}{k} - KP(x \mid n, k, \mathrm{D}^{\mathrm{binom}}_n(k))\right] = \mathrm{D}^{\mathrm{binom}}_n(k) + O(1). \quad (1)$$

According to [1], the x in (1) is a collective if $\log_2 \binom{n}{k} - K(x \mid n, k)$ is small (intuitively, this means that the complexity of x in the class of sequences in $2^{(n)}$ containing k 1s is close to maximal). The expression in square brackets in (1) is completely analogous to $\log_2 \binom{n}{k} - K(x \mid n, k)$, except for the term $\mathrm{D}^{\mathrm{binom}}_n(k)$. We can get rid of it at the expense of a certain loss of sharpness.

Corollary. *For a fixed $\epsilon > 0$,*

$$\mathrm{D}^{\mathrm{binom}}_n(k) - O(1) \le \mathrm{D}^{\mathrm{Bernoul}}(x) - \left[\log_2 \binom{n}{k} - KP(x \mid n, k)\right] \\ \le (1 + \epsilon)\, \mathrm{D}^{\mathrm{binom}}_n(k) + O(1).$$

The word "fixed" in the statement of the corollary means that $|O(1)|$ is bounded by a value that depends on ϵ. The quantities $\log_2 \binom{n}{k} - KP(x \mid n, k)$ and $\log_2 \binom{n}{k} - K(x \mid n, k)$ differ by at most

$$2 \log_2 \left(\left| \log_2 \binom{n}{k} - K(x \mid n, k) \right| + 1 \right) + O(1)$$

(we refer to this as coincidence to within $2 \log$).

Therefore, our definition of Bernoulliness adds to the requirement of nearly maximal complexity of x in the class of sequences in 2^* with the same length and the same number of 1s as x the requirement of binomiality of the number of 1s. It turns out that binomiality deficiency can be characterized in complexity-theoretic terms; this gives a complexity-theoretic characterization of Bernoulli sequences.

If $\mathfrak{A}$ is a partition (of a set into disjoint subsets), $\mathfrak{A}(k)$ denotes the element of the partition containing k.

Theorem 2. *Let $n > 0$ be an integer. Set*

$$k_s = \frac{n}{2}\left(1 - \cos\frac{s}{\sqrt{n}}\right) \quad \textit{for} \quad s = 0, 1, \ldots, \lfloor \pi\sqrt{n} \rfloor, \tag{2}$$

where $\lfloor\cdot\rfloor$ is integer part. Denote by $\mathfrak{A}$ the partition of the set $\{0, 1, \ldots, n\}$ into the subsets $[k_s, k_{s+1})$, where $s = 0, 1, \ldots, \lfloor \pi\sqrt{n} \rfloor$ (for $s = \lfloor \pi\sqrt{n} \rfloor$ we set $k_{s+1} = +\infty$). If $k \in \{0, 1, \ldots, n\}$,

$$\mathrm{D}_n^{\mathrm{binom}}(k) = \log_2 |\mathfrak{A}(k)| - KP(k \mid n, \mathfrak{A}(k)) + O(1). \tag{3}$$

The right-hand side of (3) coincides with $\log_2 |\mathfrak{A}(k)| - K(k \mid n, \mathfrak{A}(k))$ to within $2\log$. Notice the following properties of the partition $\mathfrak{A}$. The sets $\{0\}$ and $\{n\}$ are in $\mathfrak{A}$. If $k \neq 0$ and $k \neq n$,

$$|\mathfrak{A}(k)| = \sqrt{\frac{k(n-k)}{n}} \cdot 2^{O(1)}.$$

It is easy to see that

$$\sqrt{\frac{k(n-k)}{n}} = \sqrt{n\frac{k}{n}\left(1 - \frac{k}{n}\right)}$$

is an estimate of the standard deviation of the number of 1s.

The author is deeply grateful to his supervisor A. N. Kolmogorov for valuable discussions. V. V. V'yugin's and A. K. Zvonkin's comments contributed to the improvement of this note, and the author expresses his sincere gratitude to them.

References

[1] Andrei N. Kolmogorov. Logical basis for information theory and probability theory. *IEEE Transactions of Information Theory*, IT-14:662–664, 1968. Russian original: К логическим основам теории информации и теории вероятностей (published in 1969 in *Проблемы передачи информации*).

[2] Leonid A. Levin. On the notion of a random sequence. *Soviet Mathematics Doklady*, 14:1413–1416, 1973. Russian original: Л. А. Левин. О понятии случайности. *Доклады АН СССР* 212(1):548–550, 1973.

[3] Vladimir V. V'yugin. Algorithmic entropy (complexity) of finite objects and its applications to defining randomness and amount of information. *Selecta Mathematica Sovietica*, 13:357–389, 1994. Russian original: В. В. Вьюгин. Алгоритмическая энтропия (сложность) конечных объектов и ее применение к определению случайности и количества информации. *Семиотика и информатика* 16:14–43, 1981.

[4] Hartley Rogers, Jr. *Theory of Recursive Functions and Effective Computability.* McGraw-Hill, New York, 1967.

[5] Leonid A. Levin. Uniform tests of randomness. *Soviet Mathematics Doklady*, 17:337–340, 1976. Russian original: Л. А. Левин. Равномерные тесты случайности. *Доклады АН СССР* 227(1):33–35, 1976.

Scientific Council for Cybernetics
Academy of Sciences of the USSR

Received by the Board of Governors
on 6 April 1984

A Historical background

Kolmogorov's definition of Bernoulliness was part of his project of creating new mathematical foundations for applications of probability. By that time, the measure-theoretic foundations for the theory of probability clearly articulated in his *Grundbegriffe* [8] had been accepted by the community of researchers working in mathematical probability and statistics (see, e.g., [15]; an important role in the acceptance belonged to Doob's 1953 book [7]). In his approach to the foundations of applications of probability, Kolmogorov followed Richard von Mises (referring to [14] in Section I.2 of [8]). Richard von Mises suggested frequentist foundations based on the notion of a collective; his original definition was not rigorous, but two different formalizations were suggested by Abraham Wald [20] and Alonzo Church [6]. Collectives are often regarded as the first attempt to define the notion of a random sequence (roughly, a sequence that is a typical realization of a sequence of independent and identically distributed trials).

Kolmogorov regarded collectives as important only for foundations of applications of probability. He believed that there was no need to change the existing foundations of the theory of probability based on Kolmogorov's [8] axioms: "there is no need whatsoever to change the established construction of the mathematical probability theory on the basis on the general theory of measure" ([11], Section 6).

Kolmogorov's first attempt to bring von Mises's infinitary notion of collectives closer to the needs of practice was his 1963 paper [9]. After introducing his algorithmic notion of complexity in 1965 [10], Kolmogorov used it in [1] to define a finitary notion of a binary collective that was in some sense universal (being based on a universal notion of algorithmic complexity) and so immune to known examples showing the inadequacy of collectives as formalization of random sequences (such as Ville's [16] demonstration that collectives do not necessarily satisfy the conclusion of the law of the iterated logarithm). The same definition was published earlier by Per Martin-Löf ([13], Section V), Kolmogorov's PhD student.

Kolmogorov and Martin-Löf used the term "Bernoulli sequences" for their formalization of random sequences, since they considered only the case of binary random sequences, where the goal is to formalize typical realizations of repeated independent Bernoulli trials. (Kolmogorov was very much against what he regarded as premature generalizations and insisted that the simple binary case should be understood first.)

Two important features of Kolmogorov's approach to foundations of applications of probability were its finitary character ("we do not often see infinite sequences around us, do we?") and avoiding probability measures when defining random sequences. (Perhaps probability measures were to reappear at a later stage as frequencies in random sequences, as in von Mises's writings.) However, his PhD students, first of all Per Martin-Löf [13] and Leonid Levin [2, 5], were quick to bring into Kolmogorov's theory both aspects that Kolmogorov himself avoided (infinite sequences and probability measures). In terms of probability measures, Kolmogorov's preferred notion of randomness could be expressed as randomness with respect to uniform probability measures on finite sets.

B This note

Note [17] defined Bernoulliness as randomness with respect to the class of Bernoulli measures and compared the resulting notion with that of Kolmogorov [1] and Martin-Löf [13]. The latter was identified with the randomness with respect to the exchangeable distributions (4) (without mentioning them explicitly). The first main result (Theorem 1) of [17] was that the Bernoulliness deficiency decomposes into the sum of the exchangeability deficiency and the binomiality deficiency of the number of 1s, where the binomiality deficiency is defined to be the randomness with respect to the class of binomial measures. The second main result (Theorem 2) expressed the binomiality deficiency of a number k in Kolmogorov's preferred terms, as the randomness deficiency of k with respect to the uniform probability measure on a certain neighbourhood of k. Therefore, the new notion of Bernoulliness as a whole was expressed in Kolmogorov's preferred terms.

Note [17] was published in a Russian journal routinely translated into English cover-to-cover. The current translation uses the one in [17] but follows the Russian original somewhat less closely. In particular, it sets the terms being defined (such as "description method", "complexity", etc.) in italics.

C Proofs

The proofs of Theorems 1 and 2 in [17] have never been published, but they are not difficult to extract from the existing publications, such as [19] (Theorem 1) and [18] (Lemmas 1–3). This section will spell them out.

Proof of Theorem 1

Let us set, for $x \in 2^{(n)}$ and $y \in X$,

$$\mathrm{D}^{\mathrm{exch}}(x \mid y) := \log_2 \binom{n}{k} - KP(x \mid n, k, y), \tag{4}$$

where k is the number of 1s in x. This is the exchangeability deficiency of x (the randomness deficiency $\inf_P D(x \mid P; y)$ with respect to all exchangeable distributions P on $2^{(n)}$, conditioned on knowing y). In terms of $\mathrm{D}^{\mathrm{exch}}$, our goal (1) can be rewritten as

$$\mathrm{D}^{\mathrm{Bernoul}}(x) = \mathrm{D}^{\mathrm{exch}}(x \mid \mathrm{D}^{\mathrm{binom}}_n(k)) + \mathrm{D}^{\mathrm{binom}}_n(k) + O(1). \tag{5}$$

According to Theorem 1 in [19], we have

$$D(x \mid B_{n,p}) = D(k \mid \mathrm{bin}_{n,p}) + D(x \mid 2^n_k; D(k \mid \mathrm{bin}_{n,p})) + O(1) \tag{6}$$

where $B_{n,p}$ is the Bernoulli measure with parameters (n,p), $\mathrm{bin}_{n,p}$ is the binomial measure with parameters (n,p), k is the number of 1s in x, 2^n_k is the set of all sequences in $2^{(n)}$ with k 1s (identified with the uniform probability measure on this set), and $D(x \mid 2^n_k; D(k \mid \mathrm{bin}_{n,p}))$ stands for the randomness deficiency of x in 2^n_k conditioned on knowing $D(k \mid \mathrm{bin}_{n,p})$. (Namely, (6) is obtained from the equation in Theorem 1 in [19] by applying log to both sides; this is needed since the exposition in [19] is in terms of "level of impossibility" 2^{-D}, where D is randomness deficiency.)

Roughly, (5) corresponds to minimizing both sides of (6) over $p \in [0,1]$; this works because of the following theorem (which is a version of Theorem 2 in [18]).

Theorem 3. *There exists a computable point estimator $E : 2^* \to [0,1]$ such that*

$$\mathrm{D}^{\mathrm{Bernoul}}(x) = D(x \mid B_{n,E(x)}) + O(1) \tag{7}$$

and

$$\mathrm{D}^{\mathrm{binom}}_n(k) = D(k \mid \mathrm{bin}_{n,E(x)}) + O(1), \tag{8}$$

where x ranges over 2^*, n is the length of x, and k is the number of 1s in x.

An estimator E satisfying the conditions in Theorem 3 is described at the beginning of Subsection 4.1 of [18]; it depends on x only via n and k and is sometimes denoted $E_n(k)$. To define E we can, essentially, fix an element of each $\mathfrak{A}(k)/n$ (in a computable manner), and set $E_n(k)$ to the fixed element of $\mathfrak{A}(k)/n$. For agreement with [18], let us replace the partition (2) by the equivalent partition (cf. the identity $1-\cos(2\alpha) = 2\sin^2\alpha$ and Lemma 1 in [18])

$$\theta_a = n\sin^2\frac{a}{\sqrt{n}} \quad \text{for} \quad a = 0,1,\dots,\lfloor \pi\sqrt{n}/2 \rfloor.$$

It will be convenient (as in [18]) to allow a to be any number in the interval $[0, \pi\sqrt{n}/2]$.

Theorem 3 and (6) immediately imply (5):

$$\begin{aligned}\mathrm{D}^{\mathrm{exch}}&(x \mid \mathrm{D}^{\mathrm{binom}}_n(k)) + \mathrm{D}^{\mathrm{binom}}_n(k)\\ &= D(x \mid 2^n_k; D(k \mid \mathrm{bin}_{n,E(x)})) + D(k \mid \mathrm{bin}_{n,E(x)}) + O(1)\\ &\qquad = D(x \mid B_{n,E(x)}) + O(1) = \mathrm{D}^{\mathrm{Bernoul}}(x) + O(1).\end{aligned}$$

Proof of Theorem 3. Let us check (7); the proof of (8) is similar. We are required to prove

$$D(x \mid B_{n,p}) \ge D(x \mid B_{n,E(x)}) - O(1).$$

We will do this separately for the cases $|a - \hat{a}(x)| < 1$ and $|a - \hat{a}(x)| \ge 1$, where a is defined by $\theta_a = p$ and $\hat{a}$ is defined before Lemma 2 in [18].

If $|a - \hat{a}(x)| < 1$,

$$\begin{aligned}D(x \mid B_{n,p}) &= -\log_2 B_{n,p}\{x\} - KP(x \mid n,p) + O(1)\\ &\ge -\log_2 B_{n,E(x)}\{x\} - KP(x \mid n, E(x)) - O(1)\end{aligned}$$

$$= D(x \mid B_{n,E(x)}) - O(1),$$

where the inequality uses Lemma 2 in [18] (which ensures $-\log_2 B_{n,p}\{x\} \geq -\log_2 B_{n,E(x)}\{x\} - O(1)$) and $KP(x \mid n, E(x)) \geq KP(x \mid n, p) - O(1)$.

If $|a - \hat{a}(x)| \geq 1$,

$$\begin{aligned} D(x \mid B_{n,p}) &= -\log_2 B_{n,p}\{x\} - KP(x \mid n, p) + O(1) \\ &\geq -\log_2 B_{n,E(x)}\{x\} + \epsilon\,|a - \hat{a}(x)| \\ &\quad - KP(x \mid n, E(x)) - 2\log\left(|a - \hat{a}(x)| + 1\right) - O(1) \\ &\geq D(x \mid B_{n,E(x)}) - O(1), \end{aligned}$$

where $\epsilon > 0$ is a universal constant and the inequality uses Lemma 3 in [18] and the standard bound $KP(m) \leq 2\log m + O(1)$ for $m \in \{1, 2, \dots\}$. □

Proof of Theorem 2

Theorem 2 will follow from (8) in Theorem 3 and the definition of the estimator E if we show that the restriction of $\mathrm{bin}_{n,E_n(k)}$ to the subsets of $\mathfrak{A}(k)$ coincides, to within a constant factor, with the uniform probability measure on $\mathfrak{A}(k)$, where n ranges over N and k over $\{1, \dots, n\}$. By Lemma 2 and Corollary 3 in [18], it suffices to show that $\mathrm{bin}_{n,\hat{p}(k')}(k')$ coincides, to within a constant factor, with $(k'(n-k')/n)^{-1/2}$, k' ranging over $\mathfrak{A}(k)$, where $\hat{p}(k') = k'/n$ is the maximum likelihood estimate of the parameter p. It remains to apply Stirling's formula.

D Arcsine transform

The justification of the partition (2) given at the end of the note is in terms of the standard deviation of k/n. The partition itself is not motivated, but it was obtained from the differential equation

$$\frac{\mathrm{d}p}{\sqrt{p(1-p)}} = \mathrm{d}u, \tag{9}$$

p ranging over the interval $[0, 1]$. Its solution $u = u(p)$ gives a random variable which has an approximately constant variance when we plug k/n in place of p. This again follows from the variance of k/n being proportional to $p(1-p)$.

Solving (9), we obtain

$$p = \frac{1}{2}(1 - \cos u) = \sin^2 \frac{u}{2}, \quad u \in [0, \pi].$$

The inverse function, u as function of p, is a "variance-stabilizing transformation". Its standard representation is

$$u = 2 \arcsin \sqrt{p}.$$

This transformation was proposed by Zubin [22] (following Hotelling's suggestion and based on earlier work by Fisher), and it is known (with the coefficient of 2 omitted) as the *arcsine transform*. The arcsine transform is popular in sciences; not only is it widely used [12] but also widely abused [21].

Additional references

[6] Alonzo Church. On the concept of a random sequence. *Bulletin of the American Mathematical Society*, 46:130–135, 1940.

[7] Joseph L. Doob. *Stochastic Processes*. Wiley, New York, 1953.

[8] Andrei N. Kolmogorov. *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Springer, Berlin, 1933. English translation: Foundations of the Theory of Probability. Chelsea, New York, 1950.

[9] Andrei N. Kolmogorov. On tables of random numbers. *Sankhya. Indian Journal of Statistics A*, 25:369–376, 1963.

[10] Andrei N. Kolmogorov. Three approaches to the quantitative definition of information. *Problems of Information Transmission*, 1:1–7, 1965. Russian original: Три подхода к определению понятия "количество информации".

[11] Andrei N. Kolmogorov. Combinatorial foundations of information theory and the calculus of probabilities. *Russian Mathematical Surveys*, 38:29–40, 1983. Russian original: Комбинаторные основания теории информации и исчисления вероятностей.

[12] Louis Laurencelle and Denis Cousineau. Analysis of proportions using arcsine transform with any experimental design. *Frontiers in Psychology*, 13:1045436, 2023.

[13] Per Martin-Löf. The definition of random sequences. *Information and Control*, 9:602–619, 1966.

[14] Richard von Mises. *Vorlesungen aus dem Gebiete der angewandten Mathematik. I. Band. Wahrscheinlichkeitsrechnung und ihre Anwendung in der Statistik und theoretischen Physik* (Lectures on Applied Mathematics. Vol. 1. Probabilities and their Applications in Statistics and Theoretical Physics). Franz Deuticke, Leipzig and Vienna, 1931.

[15] Glenn Shafer and Vladimir Vovk. The origins and legacy of Kolmogorov's *Grundbegriffe*. The Game-Theoretic Probability and Finance project, `http://probabilityandfinance.com`, Working Paper 4, April 2013. Part of this technical report (covering the *Grundbegriffe* and the period before its publication) appeared as: The sources of Kolmogorov's *Grundbegriffe*. *Statistical Science*, 21:70–98, 2006.

[16] Jean Ville. *Etude critique de la notion de collectif*. Gauthier-Villars, Paris, 1939.

[17] Vladimir Vovk. On the concept of the Bernoulli property. *Russian Mathematical Surveys*, 41:247–248, 1986. Russian original: В. Г. Вовк. О понятии бернуллиевости. *Успехи математических наук*, 41(1):185–186, 1986.

[18] Vladimir Vovk. Learning about the parameter of the Bernoulli model. *Journal of Computer and System Sciences*, 55:96–104, 1997.

[19] Vladimir Vovk and Vladimir V. V'yugin. On the empirical validity of the Bayesian method. *Journal of the Royal Statistical Society B*, 55:253–266, 1993.

[20] Abraham Wald. Die Widerspruchfreiheit des Kollectivbegriffes der Wahrscheinlichkeitsrechnung. *Ergebnisse eines Mathematischen Kolloquiums*, 8:38–72, 1937.

[21] David I. Warton and Francis K. C. Hui. The arcsine is asinine: the analysis of proportions in ecology. *Ecology*, 92(1):3–10, 2011.

[22] Joseph Zubin. Note on a transformation function for proportions and percentages. *Applied Psychology*, 19:213–220, 1935.