Uluslararası Sosyal Araştırmalar Dergisi / The Journal of International Social Research
Cilt: 11 Sayı: 55 Şubat 2018 Volume: 11 Issue: 55 February 2018
www.sosyalarastirmalar.com Issn: 1307-9581
http://dx.doi.org/10.17719/jisr.20185537240

THE CALCULATION AND INTERPRETATION OF THE ENTROPIES OF THE TEST RESULT IN EDUCATION

Tuğba ÖZKAL YILDIZ •
Uğur ATEŞ **

Abstract

In this study, entropy that is an important subject in information theory is applied in test results in education with the aim of measurement and evaluation. Firstly, the test that is approved in terms of reliability, validity, and applicability is administered. The test is applied to 31 chosen freshmen out of 180 with simple random sampling from Vezirköprü Vocational and Technical Anatolian High School. Some descriptive statistics on the scores obtained from subscales of the test: Mathematics, Grammar, Social Studies and Sciences, are calculated and based on the probability functions of the test scores, entropy values are computed and interpreted. According to these entropies, it is sufficient to ask 10 questions so as to be informed about readiness of students who have just started Vezirkopru Vocational and Technical Anatolian High School. Consequently, it has been concluded that the same result may be obtained by measuring student levels with the fewest questions possible for each lesson the obtained information with the test of 100 questions that includes the subscales of Mathematics, Grammar, Social Studies and Sciences. The results of entropies are evidents that a teacher can reach a conclusion in a more practical way using fewer questions while doing descriptive evaluation.

Keywords: Measurement and Evaluation in Education, Descriptive Evaluation, Entropy.

1. INTRODUCTION

A tests used in the teaching-learning process comprise a vital part of the curriculum. Thanks to data obtained from the tests it is possible to determine the level of student, effectiveness of the teacher and the efficiency of teaching programs. Inferences made as a result of the comparison of the measurement obtained after the test applied to a criteria is defined as evaluation. In the light of the assessments, teaching activities, in accordance with the students' knowledge and skills, which will be implemented in subsequent periods are developed.

Examinations are used to collect quantitative information about items in a population such as classrooms. Developing an examination is a very important duty especially for a teacher. Because the exam has to be reliable, valid and applicable. In this case reliability is the degree to which an assessment tool produces stable and consistent results. Validity refers to how well a test measures what it is purported to measure. While reliability is necessary, it alone is not sufficient. For a test to be reliable, it also needs to be valid. For example, if your scale is off by 3 kg it reads your weight every day with an excess of 3 kg. The scale is reliable because it consistently reports the same weight every day, but it is not valid because it adds 3 kg to your true weight. It is not a valid measure of your weight. Applicable means applied and evaluated easy and in a short period of time. An examination that measures students' knowledge and abilities is applied for this study.

In literature, there is no similar study about application of entropy to measurement and evaluation in education. The test that measures students' readiness is dealt with in this study. Main aim of this work using entropy values to determine the number of questions in the test. The organization of this study as follows: In section two, a fundamental concepts of measurement and evaluation in education is presented. From a general point of view, concept of entropy is given in section three. In section, methods are described. The application of information theory on measurement in education which is the main topic of the study is given, and some tables and figures are constructed for the test in section five. Some final remarks are discussed in last section.

2. CONCEPTS OF MEASUREMENT AND EVALUATION IN EDUCATION

This chapter shortly defines measurement, evaluation, assessment and three forms of assessment. Measurement generally means to determine a characteristics' amount. In other words, measurement is the assignment of number to objects or events (Pedhazur and Schmelkin, 1991). Educational measurement refers

• Assoc. Prof. Dr., Dokuz Eylul University, Department of Statistics, tugba.ozkal@deu.edu.tr
**Mathematics Teacher, Didim Anatolian High School, uugurates@gmail.com

to the use of educational assessments and the analysis of data such as scores obtained from educational assessments to infer the abilities and proficiencies of students. Three qualities are required in a test; namely reliability, validity and applicability. Reliability is the degree to which an assessment tool produces stable and consistent results. Validity refers to how well a test measures what it is purported to measure. Applicable means applied and evaluated easy and short period of time. Evaluation: Procedures used to determine whether the subject (in example student) meets a preset criteria, such as qualifying for special education services. Assessment: The process of gathering information to monitor progress and make educational decisions if necessary. An assessment may include a test, but also includes methods such as observations, interviews, behavior monitoring, etc.

There are three forms of assessment:

- Diagnostic: Pre testing students to see what they know before teaching the unit. Diagnostic assessment is used to diagnose the student's strengths and weaknesses. This will help teachers identify what they need to do to help the student become stronger.
- Formative: Assessing students' strengths and weaknesses, and providing feedback during the unit. Formative assessment is the ongoing analysis of a student's needs – we must recognize when it needs to be learned, motivated, and provided with positive reinforcement in order for it to grow.
- Summative: Testing the student's knowledge at the end of teaching a unit. Summative assessment of the measuring of the student's growth at a point. The measurement tells us how much the students have grown.

3. ENTROPY

In information theory, entropy is the expected value of an event. It can be defined as measure of uncertainty associated with a random variable. Let X be a discrete random variable with a finite number of values $x_1, x_2, \ldots x_n$ and probabilities $p_i \geq 0$ for i = 1, 2,…, n and $\sum_{i=1}^{n} p_i = 1$. The entropy or marginal entropy H(X) is defined as Equation 1,

$$H(X) = -\sum_{i=1}^{n} p_i \log p_i \tag{1}$$

by Cover and Thomas (2006) and Shannon (1948). In information theory, how much uncertainty is contained in a joint system of two random variables is measured by the joint entropy. Let X and Y be random variables, then the joint entropy H(X, Y) defined in Cover and Thomas (2006) as

$$H(X,Y) = -\sum_{x=i}\sum_{y=j} p_{ij}(x, y)\log(p_{ij}(x, y)). \tag{2}$$

A quantity that measures the mutual dependence of the two variables are defined as mutual information of the two random variables. If X and Y are independent, then knowing X does not give any information about Y and the other way around, so their mutual information is zero. In the case that mutual information is positive, marginal distributions are dependent and marginal entropies exceed total entropy by an amount equal to the mutual information.. Mutual information for the random variables X and Y is given by Cover and Thomas (2006) in Equation 3.

$$I(X,Y) = H(X) + H(Y) - H(X,Y) \tag{3}$$

4. METHODS

The test of 100 questions, equally divided into four subjects; Mathematics, Grammar, Social Studies and Science, which is approved by two specialist Mathematics teachers in terms of reliability, validity and applicability is administered to 31 randomly chosen students who have just started Vezirkopru Vocational and Technical Anatolian High School at the beginning of school year 2014-2015. Each question is graded with 4 points. The results are entitled as "poor" for points from 0 to 20, "mediocre" from 20 to 40, "good" from 40 to 60, "great" from 60 to 80 and "wonderful" from 80 to 100. According to the entropy values gathered from scores, conclusions have been reached on the optimum numbers of questions for each subject in order to measure the level of other students that are assumed to have similar qualifications. The goal of

this application is to determine the level of readiness of the students based on entropy values in a practical way.

The questions representing the lessons in the survey and the probability distribution tables are created separately for each lesson mentioned, utilizing the frequency values calculated in accordance with the scores of the questions for each lesson. With the use of these probability distribution tables, Shannon entropy values are computed for Maths, Grammar, Social Studies and Science. While examining the kind of entropy values each lesson had with residence, joint probability distribution tables are constructed separately from the frequencies obtained from Residence - Maths, Residence - Grammar, Residence - Social Studies and Residence - Science scores. The joint entropy values of all lessons and residence are calculated separately from the joint probability distribution tables. Mutual information values are computed separately for each lesson and residence, utilizing the same joint probability distribution tables. All values are clarified within the scope of the information theory.

5. RESULTS

The entropy values are calculated by using the probability distributions constructed for the random variables of Maths, Grammar, Social Studies and Science. The frequencies, probabilities and entropy values of these random variables are shown in Table 1.

Table 1. Frequency and probability table for the maths, grammar, social studies and science variables

	Maths		**Grammar**		**Social Studies**		**Science**	
	f	**P**	**f**	**P**	**f**	**P**	**f**	**P**
0-20	10	0.323	6	0.194	7	0.226	17	0.548
20-40	10	0.323	7	0.226	5	0.161	10	0.323
40-60	9	0.290	7	0.226	11	0.355	3	0.097
60-80	2	0.065	7	0.226	6	0.194	1	0.032
80-100	0	0.000	4	0.129	2	0.065	0	0.000
TOTAL	31	1	31	1	31	1	31	1
ENTROPY		**1.826**		**2.294**		**2.153**		**1.488**

In addition, the bar graph of the entropy values for these random variables is presented in Figure 1.

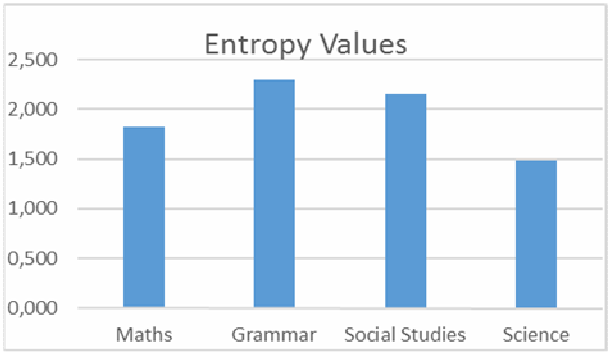

Figure 1. Entropy values for maths, grammar, social studies and science random variables

The entropy value 1.826 of Maths indicates that it is enough to ask 2 questions for Maths to have information about students' readiness. Likewise, the entropy values found for Grammar 2.294, Social Studies 2.153 and Science 1.488 also indicate that it would be enough to ask 3 questions for Grammar, 3 questions for Social Studies and 2 questions for Science.

19 out of 31 students undertaking the test are living in small villages and 12 of them are living in the county. First, these two groups are examined individually and probability distribution tables are constructed, and then entropy values are calculated. Interpretations are done according to the values. Table 2 and Table 3 show probability distribution and entropy values of Residence-Maths, Residence-Grammar, Residence-Social Studies and Residence-Science.

Table 2. Probability distribution for village and maths, grammar, social studies and science random variables

VILLAGE	Maths		Grammar		Social Studies		Science	
	f	P	f	P	f	P	f	P
0-20	10	0.526	4	0.211	4	0.211	9	0.474
20-40	9	0.4737	4	0.211	4	0.211	8	0.421
40-60	0	0	6	0.316	6	0.316	2	0.105
60-80	0	0	4	0.211	4	0.211	0	0
80-100	0	0	1	0.053	1	0.053	0	0
TOTAL	19	1	19	1	19	1	19	1
ENTROPY	0.998		2.168		2.168		1.472	

In addition, entropy values for village and Maths, Grammar, Social Studies and Science random variables are presented with bar graph in Figure 2.

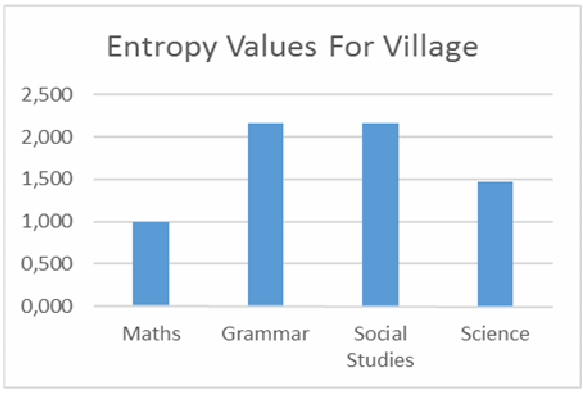

Figure 2. Entropy values for village and maths, grammar, social studies and science random variables

The entropy values calculated from the scores of students residing in villages seen in Table 2 are as follows: 0.998 for Maths, 2.168 for Grammar, 2.168 for Social Studies and 1.472 for Science. According to these entropy values, one can safely assume that 1 question for Maths, 3 for Grammar, 3 for Social Studies and 2 for Science, are adequate to gather information about the general readiness of students residing in village.

Table 3. Probability distribution for county and maths, grammar, social studies and science random variables

COUNTY	Maths		Grammar		Social Studies		Science	
	f	P	f	P	f	P	f	P
0-20	0	0	2	0.167	3	0.25	8	0.667
20-40	1	0.083	3	0.250	1	0.083	2	0.167
40-60	9	0.75	1	0.083	5	0.417	1	0.083
60-80	2	0.167	3	0.25	2	0.167	1	0.083
80-100	0	0	3	0.25	1	0.083	0	0
TOTAL	12	1	12	1	12	1	12	1
ENTROPY	1.041		1.730		2.055		1.418	

In addition, entropy values for county and Maths, Grammar, Social Studies and Science random variables are presented with bar graph in Figure 3.

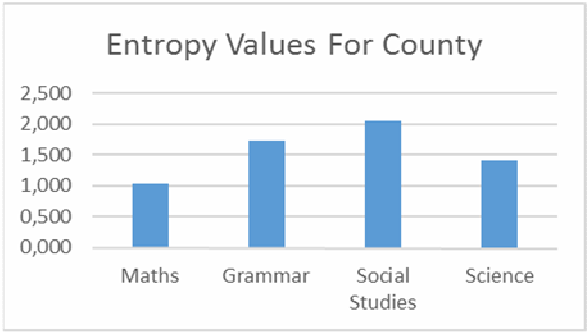

Figure 3. Entropy values for county and maths, grammar, social studies and science random variables

The entropy values calculated from the scores of students residing in counties seen in Table 4.3 are as follows: 1.041 for Maths, 1.73 for Grammar, 2.055 for Social Studies and 1.418 for Science. According to these entropy values, one can safely assume that, 2 questions for Maths, 2 for Grammar, 3 for Social Studies and 2 for Science, are adequate to gather information about the general readiness of students residing in county.

To investigate what kind of entropy values the variables of Mathematics, Grammar, Social Studies and Science considering residence, joint probability distribution tables are constructed separately from the frequencies obtained from Residence-Mathematics, Residence-Grammar, Residence-Social Studies and Residence-Science scores. Table 4 gives the joint probability distribution functions for all subscales and residence.

Table 4. Joint probability distributions of all subscales and residence

Maths

	0-20	20-40	40-60	60-80	80-100	TOTAL
Village	0.323	0.290	0.000	0.000	0.000	0.613
County	0.000	0.032	0.290	0.065	0.000	0.387
TOTAL	0.323	0.323	0.290	0.065	0.000	1.000

Grammar

	0-20	20-40	40-60	60-80	80-100	TOTAL
Village	0.129	0.129	0.194	0.129	0.032	0.613
County	0.065	0.097	0.032	0.097	0.097	0.387
TOTAL	0.194	0.226	0.226	0.226	0.129	1.000

Social Studies

	0-20	20-40	40-60	60-80	80-100	TOTAL
Village	0.129	0.129	0.194	0.129	0.032	0.613
County	0.097	0.032	0.161	0.065	0.032	0.387
TOTAL	0.226	0.161	0.355	0.194	0.065	1.000

Science

	0-20	20-40	40-60	60-80	80-100	TOTAL
Village	0.290	0.258	0.065	0.000	0.000	0.613
County	0.258	0.065	0.032	0.032	0.000	0.387
TOTAL	0.548	0.323	0.097	0.032	0.000	1.000

According to these joint probability distributions, joint entropy values are calculated as Table 5.

Table 5. Joint entropy values for all subscales and residence

Variables	Maths	Grammar	Social Studies	Science
Joint Entropy	1.977	3.155	3.087	2.356

In addition, the bar graph of joint entropy values is presented in Figure 4.

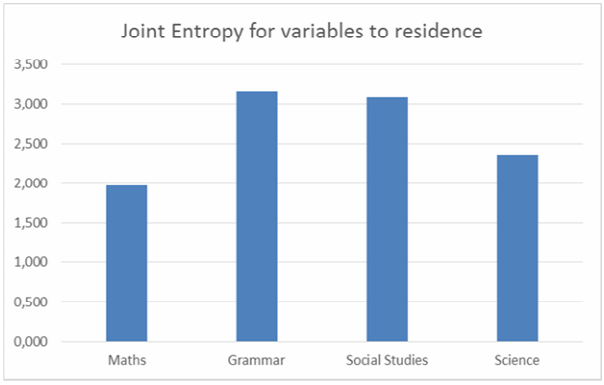

Figure 4. Joint entropy values for all subscales and residence

The result in the joint entropy seen in Table 5 H(X, Y) = 1.977 with X = Residence and Y = Maths means that on average it would require 2 questions to guess the level of both variables. The similar result is 4 questions for Grammar, 4 questions for Social Studies and 3 questions for Science.

Mutual information values are computed using the joint probability distribution calculated for residence and all subscales. The entropy value is found as 0.963 bits for residence and mutual information values calculated for Residence–Maths, Residence–Grammar, Residence–Social Studies and Residence–Science. Mutual information values are given for all variables in Table 6.

I(Maths, Residence) = H(Maths) + H(Residence) – H(Maths, Residence)
I(Maths, Residence) = 1.826 + 0.963 – 1.977 = 0.812 bits.

I(Grammar, Residence) = H(Grammar) + H(Residence) – H(Grammar, Residence)
I(Grammar, Residence) = 2.294 + 0.963 – 3.155 = 0.102 bits.

I(Social Studies, Residence) = H(Social Studies) + H(Residence) – H(Social Studies, Residence)
I(Social Studies, Residence) = 2.153 + 0.963 – 3.087 = 0.029 bits.

I(Science, Residence) = H(Science) + H(Residence) – H(Science, Residence)
I(Science, Residence) = 1.488 + 0.963 – 2.356 = 0.095 bits.

Table 6. Mutual information for variables and residence

Variables	Maths	Grammar	Social Studies	Science
Mutual Information	0.812	0.102	0.029	0.095

The interpretation of the calculated mutual information values seen in Table 6 is as follows. The mutual information value calculated for the Maths-Residence is 0.812 bits. Then the variables Maths and Residence seem to have some information in common. Knowing the residence of any student provides information about the level of the students' Maths score. The information values calculated for the Grammar-Residence, Social Studies-Residence and Science-Residence are 0.102, 0.029 and 0.095 bits, respectively. As a result there seem to have a little information in common between these variables.

6. CONCLUSIONS

Detailed interpretation of the calculated entropy values seen in Table 1, Table 2, Table 3, Table 5 and Table 6 are given in after each table. As in Table 1, it is sufficient to ask 10 questions so as to be informed about readiness of students who have just started Vezirkopru Vocational and Technical Anatolian High

School. The entropy values calculated from the scores of students residing in villages and county are seen in Table 2 and Table 3 respectively. According to these entropy values, one can safely assume that 9 questions in total, are adequate to gather information about the general readiness of students residing in both village and county. Although for the students in both groups 9 questions in total are sufficient, questions needed for each lesson vary. As it is seen in the joint entropy values in Table 5, it is enough to ask 13 questions in total. Thus, asking 13 questions is a healthier way to determine students' readiness. Finally, from the Table 6, the mutual information measures tell us that residence plays an important role for students' scores, mostly in Maths.

Consequently, with the acquired entropy values it has been concluded that the same result may be obtained by measuring student levels with the fewest questions possible for each lesson. The joint and marjinal entropy values have been calculated based on residence of students. Thus, it is evident that a teacher can reach a conclusion in a more practical way using fewer questions while doing descriptive evaluation.

REFERENCES

Cover, M. Thomas and Thomas, A. Joy (2006). *Elements of Information Theory (2nd Ed.)*. New Jersey: John Wiley & Sons.
Pedhazur, J. Elazar and Schmelkin, P. Liora (1991). *Measurement, design, and analysis: an integrated approach* (1st Ed.) . pp. 15-29. Hillsdale, NJ: Lawrence Erlbaum Associates.
Shannon, E.Claude (1948). A Mathematical theory of communication. *The Bell System Technical Journal*, Vol. 27, pp. 379-423.