On independent sets in purely atomic probability spaces with geometric distribution

Eugen J. Ionascu and Alin A. Stancu

Department of Mathematics, Columbus State University
Columbus, GA 31907, USA
ionascu_eugen@colstate.edu, *stancu_alin1@colstate.edu*

August 12th, 2008

Abstract

We are interested in constructing concrete independent events in purely atomic probability spaces with geometric distribution. Among other facts we prove that there are uncountable many sequences of independent events.

1 Introduction

Let us assume a fixed ratio r is given, $r \in (0,1)$. In what follows we will work with the discrete probability space $\mathbb{N}_0 = \{0,1,2,3,...\}$ and the usual geometric probability on $\mathcal{A}$ (all subsets of $\mathbb{N}_0$) defined by

$$P_r(E) := \frac{1-r}{r} \sum_{k \in E \setminus \{0\}} r^k \ for\ every\ set\ E \in \mathcal{A}.$$

We are interested to study the class of independent sets in this probability space. We are going to follow [2] and define:

Mathematics Subject Classification: Primary 60A10
Key Words: Independence, purely atomic probability, geometric distribution

$A, B \in \Omega$ *are called independent if* $P(A \cap B) = P(A)P(B)$.

With this definition for every $E \subset \Omega$, Ω and E are independent and $\emptyset$ and E are also independent. These are clearly trivial examples. Three or more subsets of Ω, A_1, ..., A_n are called *mutually independent* or simply *independent* if for every choice of k ($n \geq k \geq 2$) such sets, say A_{i_1},...,A_{i_k}, we have

$$P(\bigcap_{j=1}^{k} A_{i_j}) = \prod_{j=1}^{k} P(A_{i_j}). \tag{1}$$

So, for n ($n \geq 2$) independent sets one needs to have $2^n - n - 1$ relations as in (1) to be satisfied. An infinite family of subsets is called independent if each finite collection of these subsets is independent. Events are called trivial if their probability is 0 or 1.

If $n \in \mathbb{N}$ then $\Omega(n)$ usually denotes the number of primes dividing n counting their multiplicities (see [8]). In [1] and [6], independent families of events have been studied for finite probability spaces with uniform distribution. Eisenberg and Ghosh [6] show that the number of nontrivial independent events in such spaces cannot be more than $\Omega(m)$ where m is the cardinality of the space. This result should be seen in view of the known fact (see Problem 50, Section 4.1 in [7]) that if A_1, A_2, ..., A_n are independent non-trivial events of a sample space X then $|X| \geq 2^n$. One can observe that in general $\Omega(m)$ is considerably smaller than $\log_2 m$. It is worth mentioning that according to [5] the first paper to deal to this problem in uniform finite probability spaces is [9]. In their paper, Shiflett and Shultz [9] raise the question of the existence of spaces with no non-trivial independent pairs, called *dependent probability spaces*. A space containing non-trivial independent events is called *independent*. For uniform distributed probability spaces X, as a result of the work in [6] and [1], X is dependent if $|X|$ a prime number and independent if $|X|$ is composite. For denumerable sets X one can see the construction given in [5] or look at the Example 1.1 in [10]. For our spaces, the Example 1.1 does not apply and in fact, we will construct explicitly lots of independent sets.

For every $n \in \mathbb{N}$ one can consider the following space of geometric probability distribution, denoted here by $\mathcal{G}_n := ([n], \mathcal{P}([n]), P)$ where $P(k) = q^k$ with $k \in [n] := \{1, 2, 3, ..., n\}$ and of course q is the positive solution of the equation

$$\sum_{k=1}^{n} q^k = 1.$$

This space is independent for every $n \geq 4$ with n composite. Indeed, if $n = st$ with $s, t \in \mathbb{N}$ $s, t \geq 2$ on can check that the sets $A := \{1, 2, 3, ..., s\}$, $B := \{1, s+1, 2s+1..., (t-1)s+1\}$ represent non-trivial independent events. To match the uniform distribution situation, it would be interesting if $\mathcal{G}_n$ was a dependent space for every n prime.

The class of independent sets is important in probability theory for various reasons. Philosophically speaking, the concept of independence is at the heart of the axiomatic system of modern probability theory introduced by A. N. Kolmogorov in 1933. More recently, it was shown in [3] that two probability measures on the same space which have the same independent (pairs of) events must be equal if at least one of them is atomless. This was in fact a result of A. P. Yurachkivsky from 1989 as the same authors of [3] point out in the addendum to their paper that appeared in [4].

On the other hand, Szekely and Mori [10] show that if the probability space is atomic then there may be no independent sets or one may have a sequence of such sets. The following result that appeared in [10] is a sufficient condition for the existence of a sequence of independent events in the probability space.

Theorem 1. *If the range of a purely atomic probability measure contains and interval of the form* $[0, \epsilon)$ *for some* $\epsilon > 0$ *then there are infinitely many independent sets in the underlying probability space.*

Let us observe that, if $r = 1/2$ the probability space $(\mathbb{N}_0, \mathcal{A}, P_{1/2})$ does satisfy the hypothesis of the above theorem with $\epsilon = 1$ because every number in $[0, 1]$ has a representation in base 2. On the other hand if, let us say $r = 1/3$, then the range of $P_{1/3}$ is the usual Cantor set which has Lebesgue measure zero, so Theorem 1 does not apply to $(\mathbb{N}_0, \mathcal{A}, P_{1/3})$. However, we will show that there are uncountably many pairs of sets that are independent in $(\mathbb{N}_0, \mathcal{A}, P_r)$ for every $0 < r < 1$ (these sets do not depend of r).

2 Independent pairs of events for denumerable spaces

The first result we would like to include is in fact a characterization, under some restrictions of r, of all pairs of independent events (A, B), in which one of them, say B, is fixed and of a certain form. This will show in particular that there are uncountably many such pairs. In order to state this theorem we need to start with a preliminary ingredient.

Lemma 1. *For* $m \geq 1$, *consider the function given by*

$$f(x) = (2x - 1)(1 + x^m) - x^m \text{ for all } x \in [0, 1].$$

The function f is strictly increasing and it has unique zero in $[0,1]$ denoted by t_m. Moreover, for all m we have $t_m > 1/2$, the sequence $\{t_m\}$ is decreasing and

$$\lim_{m\to\infty} t_m = \frac{1}{2}.$$

Having t_m defined as above we can state our first theorem.

Theorem 2. *For every natural number $n \geq 2$, we define the events $E := \{0, n-1\}$ and*

$$B := \{\underbrace{1,2,...,n-1}_{n-1}, \underbrace{2n-1, 2n, ..., 3n-3}_{n-1}, \\ \underbrace{4n-3, 4n-2, ..., 5n-5}_{n-1}, ...\}. \tag{2}$$

Also, for $T \subset B$ an arbitrary nonempty subset we set $A := E + T$ with the usual definition of addition of two sets in a semigroup. Then A and B are independent events in $(\mathbb{N}_0, \mathcal{A}, P_r)$.

Conversely, if $r < t_m$ (where $m = n-1$ and t_m as in Lemma 1), B is given as in (2) and A forms an independent pair with B, then A must be of the above form, i.e. $A = E + T$ for some $T \subset B$.

Proof of Lemma 1. The function f has derivative $f'(x) = 2(1+x^m) - 2m(1-x)x^{m-1}$, $x \in (0,1]$. For $m \geq 2$, using the Geometric-Arithmetic Mean inequality we have

$$(m-1)(1-x)x^{m-1} \leq \left[\frac{(m-1)(1-x) + \overbrace{x + x + ... + x}^{m-1}}{m}\right]^m = \left(\frac{m-1}{m}\right)^m$$

and so $m(1-x)x^{m-1} \leq (\frac{m-1}{m})^{m-1} \leq 1$ which implies $m(1-x)x^{m-1} \leq 1$. This last inequality is true for $m = 1$ too. This implies that

$$f'(x) = 2(1+x^m) - 2m(1-x)x^{m-1} \geq 2x^m > 0$$

for all $x \in (0,1]$. Therefore the function f is strictly increasing and because $f(1/2) = -\frac{1}{2^m} < 0$ and $f(1) = 1 > 0$, by the Intermediate Values Theorem there must be an unique solution $x = t_m$, of the equation $f(x) = 0$ in the interval $(1/2, 1)$. Because $f(t_{m-1}) = \left(\frac{1-t_{m-1}}{1+t_{m-1}^{m-1}}\right) t_{m-1}^{m-1} > 0$ we see that

$t_m < t_{m-1}$ for all $m \geq 2$. Since $(2t_m - 1)(1 + t_m^m) = t_m^m$ we can let m go to infinity in this equality and obtain $t_m \to 1/2$. ■

Using Maple, we got some numerical values for the sequence t_m: $t_1 = \frac{1}{\sqrt{2}} \approx 0.707$, $t_2 \approx 0.648$, $t_3 \approx 0.583$, $t_4 \approx 0.539$ and for instance $t_{10} \approx 0.5005$.

Proof of Theorem 2. First let us check that $E_1 = E + 1 = \{1, n\}$ and B are independent. Since $E_1 \cap B = \{1\}$, $P(\{1\}) = \frac{1-r}{r} r = 1 - r$ and $P_r(E_1) = \frac{1-r}{r}(r + r^n) = (1-r)(1 + r^{n-1})$, we have to show that $P_r(B) = \frac{1}{1+r^{n-1}}$. We have

$$P_r(B) = \frac{1-r}{r}\left(\sum_{j=1}^{m} r^j\right)\left(\sum_{i=0}^{\infty} r^{2mi}\right) = \frac{r - r^{m+1}}{r}\frac{1}{1-r^{2m}} = \frac{1}{1+r^m}$$

which is what we needed. Now, suppose $b \in B$ and consider $E_b = E + b = \{b, b+n-1\}$. We notice that by the definition of B, the intersection $B \cap E_b$ is $\{b\}$. Hence, $P_r(B \cap E_b) = \frac{1-r}{r} r^b = (1-r)r^c$ (with $c = b - 1$) and

$$P_r(B)P_r(E_b) = \frac{1}{1+r^m}\frac{1-r}{r}\left(r^b + r^{b+m}\right) = (1-r)r^c.$$

Hence, B and E_b are independent for every $b \in B$.

> *Next we would like to observe that if (F_1, B) and (F_2, B) are independent pairs of events and $F_1 \cap F_2 = \emptyset$, then $F_1 \cup F_2$ and B are independent events as well.*

Indeed, by the given assumption we can write

$$P_r(B \cap (F_1 \cup F_2)) = P_r((B \cap F_1) \cup (B \cap F_2)) = P_r(B \cap F_1) + P_r(B \cap F_2) =$$

$$P_r(B)P_r(F_1) + P_r(B)P_r(F_2) = P_r(B)(P_r(F_1) + P_r(F_2)) = P_r(B)P_r(F_1 \cup F_2).$$

In fact, the above statement can be generalized to a sequence of sets F_k which are pairwise disjoint, due to the fact that P_r is a genuine finite measure and so it is continuous (from below and above). Then if $T \subset B$ is nonempty, $A = E + T = \bigcup_{b \in B} E_b$ is countable union and since $E_b \cap E_{b'} = \emptyset$ for all $b, b' \in B$ ($b \neq b'$) the above observation can be applied to $\{E_b\}_{b \in T}$. So, we get that B and A are independent.

For the converse, we need the following lemma.

Lemma 2. *If $L \subset \mathbb{N}_0 \setminus B$ and the smallest element of L is $s = (2i-1)m+j$, where $i, j \in \mathbb{N}$, $j \leq m$, then*

$$P_r(L) \leq r^{s-1} - \frac{r^{2im}}{1+r^m}.$$

Proof of Lemma 2 Indeed, we have

$$P_r(L) \leq \tfrac{1-r}{r}[(r^s + r^{s+1} + ... + r^{2im}) + (r^{(2i+1)m+1} + ...)] =$$

$$r^{s-1} - r^{2im} + r^{2im}P_r(\Omega \setminus B) = r^{s-1} - r^{2im} + r^{2im}(1 - \tfrac{1}{1+r^m}) = r^{s-1} - \tfrac{r^{2im}}{1+r^m}.$$

■

So, let us assume that $r < t_m$, B is as in (2) and A is independent of B. We let T be the intersection of A and B and we put $\alpha := P_r(T)/P_r(B)$. Also, define $A' := T + \{0, n-1\}$, $L = A \setminus A'$ and $L' = A' \setminus A$. We have clearly $L, L' \subset \Omega \setminus B$. By the first part of our theorem $P_r(A') = \alpha$. Because A and B are independent $P_r(A)$ must be equal to α as well. Hence $P_r(A) = P_r(A')$ which attracts

$$\sum_{k \in L'} r^k = \sum_{k \in L} r^k \Leftrightarrow \sum_{k \in L \cup L'} r^k = 2 \sum_{k \in L'} r^k. \tag{3}$$

From (3), it is clear that $L' = \emptyset$ if an only if $L = \emptyset$ and so if L' is empty then $A = A'$, which is what we need in order to conclude our proof. By way of contradiction, suppose $L' \neq \emptyset$ (or equivalently $L \neq \emptyset$) we can assume without loss of generality that L' contains the smallest number of $L' \cup L$, say s which is written as in Lemma 2. Thus from equality (3) we have $P_r(L \cup L') \geq 2P_r(L')$ and then by Lemma 2 we get

$$r^{s-1} - \frac{r^{2im}}{1+r^m} \geq 2(1-r)r^{s-1} \Leftrightarrow 2r \geq 1 + \frac{r^{2im+1-s}}{1+r^m} \Leftrightarrow 2r \geq 1 + \frac{r^{n-j}}{1+r^m}.$$

Therefore for every n and $1 \leq j \leq m$,

$$2r \geq 1 + \frac{r^{n-j}}{1+r^m} \geq 1 + \frac{r^m}{1+r^m} \Rightarrow f(r) = (2r-1)(1+r^m) - r^m \geq 0.$$

By Lemma 1 we see that $r \geq t_m$ which is a contradiction. It remains that L and L' must be empty and so $A = A'$. ■

In the previous theorem, since T was an arbitrary subset of an infinite set we obtain an uncountable family of pairs of independent sets.

Remark 1: If $r = \sqrt{\frac{1}{\phi}}$ where ϕ stands for the classical notation of the golden ratio (i.e. $\phi = \frac{\sqrt{5}+1}{2}$), $n = 2$, $B = \{1, 3, 5, 7, ...\}$ as in (2), and $A = \{1, 4, 6\}$, then one can check that $P_r(B) = \frac{1}{1+r}$, $P_r(A \cap B) = 1 - r$, $P_r(A) = (1 - r)(1 + r^3 + r^5)$. So the equality $P_r(A \cap B) = P_r(A)P_r(B)$ is equivalent to $1 + r = 1 + r^3 + r^5$ which is the same as $r^4 + r^2 - 1 = 0$. One can easily see that this last equation is satisfied by $r = \sqrt{\frac{1}{\phi}}$. Hence A and B are independent but clearly A is not a translation of $\{0, 1\}$ with a subset of B. Therefore the converse part in Theorem 1 cannot be extended to numbers $r \geq t_m$ such as $r = \sqrt{\frac{1}{\phi}}$. In fact, we believe that the constants t_m are sharp, in the sense that for all $r > t_m$ the converse part is false, but an argument for showing this is beyond the scope of this paper.

Remark 2: Another family of independent events which seems to have no connection with the ones constructed so far is given by $A = \{1, 2, 3, 4, ..., n-1, n\}$ and $B = \{n, 2n, 3n, ...\}$, with $n \in \mathbb{N}$. A natural question arises as a result of this wealth of independent events: can one characterize all pairs (A, B) which are independent regardless the value of the parameter r?

3 Three independent events

The next theorem deals with the situation in which two sets as in the construction of Theorem 2 form with B given by (2), a triple of independent sets.

Let us observe that if A_1, A_2, and B are mutually independent then by Theorem 2 (at least if $r \in (0, t_m)$), A_1 and A_2 must be given by $A_i = T_i + E$ with $T_i \subset B$, $i = 1, 2$. Therefore $A_1 \cap A_2 = (T_1 \cap T_2) + E$.

Also, we note that $P_r(A_i) = P_r(T_i)(1 + r^{n-1})$, $i = 1, 2$, and $P_r(A_1 \cap A_2) = P_r(T_1 \cap T_2)(1 + r^{n-1})$. This means that the equality $P_r(A_1 \cap A_2) = P_r(A_1)P(A_2)$ is equivalent to

$$P_r(T_1 \cap T_2) = P_r(T_1)P_r(T_2)(1 + r^{n-1}). \tag{4}$$

On the other hand the condition $P_r(A_1 \cap A_2 \cap B) = P_r(A_1)P_r(A_2)P(B)$ reduces to

$$P_r(T_1 \cap T_2) = P_r(T_1)P_r(T_2)(1 + r^{n-1})^2 P_r(B),$$

which is the same as (4). So, three sets A_1, A_2 and B are independent if and only if (4) is satisfied. Let us notice that the condition (4) may be interpreted as a conditional probability independence relation:

$$P_r(T_1 \cap T_2|B) = P_r(T_1|B)P_r(T_2|B). \tag{5}$$

At this point the construction we have in Theorem 2 can be repeated. As a result, regardless of what r is, we obtain an uncountable family of there events which are mutually independent in $(\mathbb{N}_0, \mathcal{A}, P_r)$.

Theorem 3. *For a fixed $n \geq 3$, we consider B as in (2), and pick $b \in \{2, ..., n-1\}$ such that $2(b-1)$ divides $m = n-1$ $(m = 2(b-1)k)$. For $F := \{0, b-1\}$, we let*

$$\begin{aligned} B_1' := \{&\underbrace{1, 2, ..., b-1}_{b-1}, \underbrace{2b-1, 2b, ..., 3b-3}_{b-1}, \underbrace{4b-3, 4b-2, ..., 5b-5}_{b-1}, \\ &..., \underbrace{(2k-2)(b-1)+1, ..., (2k-1)(b-1)}_{b-1}\}, \end{aligned} \tag{6}$$

$$B_1 := B_1' \cup (B_1' + 2m) \cup (B_1' + 4m) \cup (B_1' + 6m) \cup ...$$

and T a subset of B_1. Then $T_1 := F + T$ and B_1 are independent sets relative to the induced probability measure on B. Moreover, $A_1 := T_1 + \{0, n-1\}$, $A_2 := B_1 + \{0, n-1\}$ and B form a triple of mutually independent sets in $(\mathbb{N}_0, \mathcal{A}, P_r)$ for all r.

Proof. The second part of the theorem follows from the considerations we made before the theorem and from the first part. To show the first part we need to check (4) for T_1 and $T_2 = B_1$. Let us remember that

$$B = \{\underbrace{1, 2, ..., n-1}_{n-1}, \underbrace{2n-1, 2n, ..., 3n-3}_{n-1},$$

$$\underbrace{4n-3, 4n-2, ..., 5n-5}_{n-1}, ...\}, \text{ and } P_r(B) = \tfrac{1}{1+r^m}.$$

We observe that $B_1' \subset \{1, 2, ..., n-1\}$ and so $B_1 \subset B$. Let us first take into consideration the case $T = \{1\}$. Since $T_1 = \{1, b\}$ we get $T_1 \cap T_2 = \{1\}$, $P_r(T_1) = (1-r)(1+r^{b-1})$, and

$$P_r(B_1) = P_r(B_1')(1 + r^{2m} + r^{4m} + r^{6m} + ...) = \frac{P_r(B_1')}{1 - r^{2m}}.$$

So, it remains to calculate $P_r(B_1')$:

$P_r(B_1') = \frac{1-r}{r}(r + r^2 + ...r^{b-1})(1 + r^{2(b-1)} + r^{4(b-1)} + ... + r^{2(k-1)(b-1)})$

$$= (1 - r^{b-1})\frac{1 - r^{2k(b-1)}}{1 - r^{2(b-1)}} = \frac{1 - r^m}{1 + r^{b-1}} \Rightarrow P_r(B_1) = \frac{1}{(1 + r^{b-1})(1 + r^m).}$$

This shows that (4) is satisfied. In the general case, i.e. T an arbitrary subset of B_1, we proceed as in the proof of Theorem 1. ■

4 Uncountable sequences of independent events

In [10], Szekely and Mori give an example of an infinite sequence of independent sets in $(\mathbb{N}_0, \mathcal{A}, P_{1/2})$. Given an infinite sequence of independent sets $\{A_n\}_n$ we may assume that $P_r(A_k) \leq \frac{1}{2}$ and so by Proposition 1.1 in [10] we must have

$$\sum_{k=1}^{\infty} P_r(A_k) < \infty.$$

Let us observe that Theorem 2 can be applied to a different space now that can be constructed within B given by (2)in terms of classes: $\widehat{\mathbb{N}}_0 = \{\hat{0}, \hat{1}, \hat{2}, ...\}$ where $\hat{0} = \emptyset$, $\hat{1} := \{1, 2, ..., n-1\}$, $\hat{2} := \{2n-1, 2n, ..., 3n-3\}$, $\hat{3} := \{4n-3, 4n-2, ..., 5n-5\}$, ..., and the probability on this space is the conditional probability as subsets of B.

Hence for $k \in \mathbb{N}$, one can check that

$$P(\hat{k}) = \frac{1 - r^{2m}}{r^{2m}} r^{2km}, \ \textit{with } m = n - 1.$$

This shows that this space is isomorphic to $(\mathbb{N}_0, \mathcal{A}, P_s)$ with $s = r^{2m}$.

One can check by induction the following proposition.

Proposition 1. *Let $n \in \mathbb{N}$, $n \geq 2$. If A_1,...,A_n are independent in $\widehat{\mathbb{N}}_0$ then $A_1 + T$, $A_2 + T$,..., $A_n + T$ and B are indepenedent in $(\mathbb{N}_0, \mathcal{A}, P_r)$.*

This construction can be then iterated indefinitely giving rise of a sequence B, B_1, B_2,..., which is going to be independent and its construction is in terms of a sequence $(n, n_1, n_2, ...)$ with $n_k \geq 2$. As a result, we have a countable way of constructing sequences of independent sets. This construction coincides with the one in [10] if $n_k = 2$ for all $k \in \mathbb{N}$.

References

[1] Y. M. Baryshnikov and B. Eisenberg, *Independent events and independent experiments*, Proc. Amer. Math. Soc., vol. 118, No. 2, 1993, pp. 615-617

[2] P. Billingsley, *Probability and Measure*, 3rd ed. J. Wiley & Sons, New York, 1995

[3] Z. Chen, H. Rubin and R. A. Vitale, *Independence and determination of probabilities*, Proc. Amer. Math. Soc. **125**, No. 12, (1997)

[4] Z. Chen, H. Rubin and R. A. Vitale, *Addendum to "Independence and determination of probabilities"*, Proc. Amer. Math. Soc. **129**, No 9, (2001)

[5] W. Edwards, R. Shiflett, H. Shultz, *Dependent probability spaces*, Collega Mathematics Journal **39**, No. 3, (2008), pp. 221-226

[6] B. Eisenberg and B. K. Ghosh, *Independent events in a discrete uniform probability space*, Amer. Statist. **41** (1987), 52-56

[7] C. M. Grinstead and J. L. Snell, *Introduction to Probability*, AMS, 1997, 510 pp

[8] I. Niven, H. S. Zuckerman and H.L. Montgomery, *An introduction to the theory of numbers*, 5th ed. J. Wiley & Sons, New York, 1991

[9] R. C. Shiflett and H. S. Shultz, *An approach to independent sets*, Mathematical Spectrum **12** (1997/80), pp. 11-16.

[10] G. J. Skekely and T. F. Mori, *Independence and atoms*, Proc. Amer. Math. Soc. **130**, No 1, pp 213-216, 2001