The Fraction of Subspaces of $\mathrm{GF}(q)^n$ with a Specified Number of Minimal Weight Vectors is Asymptotically Poisson

Edward A. Bender
Center for Communications Research
4320 Westerra Court
San Diego, CA 92121, USA
ed@ccrwest.org

E. Rodney Canfield
Department of Computer Science
University of Georgia
Athens, GA 30602, USA
erc@cs.uga.edu

Submitted: August 30, 1996; Accepted: November 27, 1996

Abstract

The *weight* of a vector in the finite vector space $\mathrm{GF}(q)^n$ is the number of nonzero components it contains. We show that for a certain range of parameters (n, j, k, w) the number of k-dimensional subspaces having $j(q-1)$ vectors of minimum weight w has asymptotically a Poisson distribution with parameter $\lambda = \binom{n}{w}(q-1)^{w-1}q^{k-n}$. As the Poisson parameter grows, the distribution becomes normal.

AMS-MOS Subject Classification (1990). Primary: 05A16 Secondary: 05A15, 11T99

1. Introduction

Almost all the familiar concepts of linear algebra, such as dimension and linear independence, are valid without regard to the characteristic of the underlying field. An example of a characteristic-dependent result is that a nonzero vector cannot be orthogonal to itself; researchers accustomed to real vector spaces must modify their "intuition" on this point when entering the realm of finite fields.

Let q be a prime power, fixed for the remainder of the paper, and $\mathrm{GF}(q)$ be the finite field with q elements. Because the underlying field is finite, there are many counting problems associated with fundamental concepts of linear algebra; for example, how many subspaces of dimension k are there in the vector space $\mathrm{GF}(q)^n$? The answer is often denoted $\begin{bmatrix} n \\ k \end{bmatrix}_q$, and we have

$$\begin{bmatrix} n \\ k \end{bmatrix}_q = \frac{(1-q^n)(1-q^{n-1})\cdots(1-q^{n-k+1})}{(1-q^k)(1-q^{k-1})\cdots(1-q)},$$

the Gaussian polynomial. The reader may consult [2] for an introduction to the subject.

Define the *weight* of a vector $\mathbf{v}$ in $\mathrm{GF}(q)^n$ to be the number of nonzero coordinates in $\mathbf{v}$. The interaction of weight with familiar concepts of linear algebra yields more and harder counting problems. Consider the n vectors of weight 1 in $\mathrm{GF}(2)^n$; how many vector spaces of dimension k do they span ? The easy answer is the well known binomial coefficient $\binom{n}{k}$. Now consider the $\binom{n}{2}$ vectors of weight 2 in $\mathrm{GF}(2)^n$; how many vector spaces of dimension k do they span ? More thought is needed this time, but again the answer is a classical array from combinatorics, the Stirling numbers of the second kind $S(n, n-k)$. If we ask the same question for weight 3 or higher, no simple answer is known and the numbers cannot be computed easily. However, that familiar properties of $\binom{n}{k}$ and $S(n,k)$ persist in the higher weight version is part of a sweeping conjecture that the Whitney numbers of the second kind for any geometric lattice are log-concave. [1, p.141]

Extend the notion of weight to subspaces by saying that a subspace $V \subseteq \mathrm{GF}(q)^n$ has weight w if w is the minimum weight of all nonzero vectors in V. A natural problem is to describe how the $\begin{bmatrix} n \\ k \end{bmatrix}_q$ subspaces of dimension k are distributed by weight. We cannot give a definitive solution to this question, but using asymptotic methods, we can gain some insight into the problem. The number of weight w vectors in a vector space over $\mathrm{GF}(q)$ is a multiple of $q-1$ since multiplication by a nonzero scalar preserves weight. Let $p(j; n, k, w)$ be the fraction of k-dimensional subspaces of $\mathrm{GF}(q)^n$ containing $j(q-1)$ vectors of weight w, and no nonzero vector of weight less than w. Masol [3] showed that

$$p(0; n, k, 1) - e^{-\lambda} \to 0 \text{ uniformly as } n \to \infty \text{ where } \lambda = nq^{k-n}.$$

We extend this result as follows:

Theorem 1. *Fix a prime power q, positive constants $b \le \frac{1}{2}$ and $B < 1-b$, and a function $\mu(n) = o(1)$. Then, uniformly for j, k, w satisfying*

$$1 \;\le\; w \;\le\; \mu(n)n^b/\log_q n \quad \text{and} \quad \lambda \;\stackrel{\text{def}}{\equiv}\; \binom{n}{w}(q-1)^{w-1}q^{k-n} \;\le\; B\log_q n, \tag{1}$$

we have

$$p(j;n,k,w) - \lambda^j e^{-\lambda}/j! \to 0 \quad \text{as} \quad n \to \infty.$$

Theorem 2. *Fix a prime power q, positive constants $b \le \frac{1}{2}$ and $B < 1-b$, and a function $\mu(n) = o((\log n)^{-1/4})$. When (1) holds,*

$$\sqrt{2\pi\lambda}\; p(j;n,k,w) - e^{-(j-\lambda)^2/2\lambda} \to 0 \quad \text{uniformly as} \quad \lambda \to \infty.$$

We remind the reader of the meaning of uniformity. A function $f : N^\ell \to N$ goes to 0 as $n \to \infty$, uniformly over $A \subseteq N^\ell$, provided

$$\sup_{a \in A_n} |f(n,a)| \to 0 \;\text{ as } n \to \infty, \quad \text{where} \quad A_n \;=\; \{a \in N^{\ell-1} : (n,a) \in A\}.$$

Since $(q-1)\lambda/q^k$ is the the probability that a randomly chosen vector has weight w, the distribution of $(q-1)$-tuples of weight w vectors in k-dimensional subspaces is asymptotically the same as the distribution of weight w vectors in random samples of $\frac{q^k-1}{q-1}$ (nonzero) vectors: They are asymptotically Poisson with parameter λ. A point in projective space is the $(q-1)$-tuple of scalar multiples of a nonzero point in $\mathrm{GF}(q)^n$. Thus, in a projective n-space over $\mathrm{GF}(q)$, the previous observation states that the distribution of weight w points is asymptotically the same among random sets of $\frac{q^k-1}{q-1}$ points and among random $(k-1)$-dimensional projective subspaces, namely Poisson with parameter λ.

2. Proof of the theorems

We will find it convenient to work with the somewhat larger

$$\lambda \;\stackrel{\text{def}}{\equiv}\; \frac{(q-1)^{w-1}n^w}{q^{n-k}w!} \;\le\; B\log_q n, \tag{2}$$

rather than the definition in (1). Since the ratio of the two versions of λ is

$$\frac{\binom{n}{w}w!}{n^w} \;=\; (1-O(w/n))^w \;=\; \exp(O(w^2/n)) \;=\; \exp(o(1/(\log n)^2)),$$

it is easily seen that the ratio of the two versions of λ tends to 1 and also that the theorems for either version of λ imply the theorems for the other version. Since the ratio of the two versions of λ approach 1, the inequality in (2) follows from that in (1) by replacing B with the average of B and $1-b$.

Let $d = n-k$, $\epsilon = (1-b-B)/2$, and $J = (B+\epsilon)\log_q n$. Making $\mu(n)$ larger if necessary, we may assume

$$\mu(n) \ \geq \ \max(n^{-b}, e^{-\epsilon^2 \log_q n/4}). \tag{3}$$

Our proof consists of five parts:

(a) Eliminating small k.

(b) Some easy estimates.

(c) Conversion of the problem to the study of $d \times n$ matrices over $\mathrm{GF}(q)$.

(d) An estimate for $j \leq J$.

(e) Completion of the proof.

Although we will not explicitly state it, all estimates in $o(\cdots)$ and $O(\cdots)$ as well as all estimates of the form $\cdots \to \cdots$ are uniform.

(a) Eliminating small k. Suppose Theorem 1 holds for $k = k_0 \stackrel{\text{def}}{\equiv} \lceil n/2 \rceil$. We will deduce that it holds for $k < k_0$. From the bound on w in (1), it follows that $\lambda \to 0$ whenever $k \leq k_0$. Since Theorem 1 is equivalent to $p(0; n, k, w) = 1 - o(1)$ when $\lambda \to 0$, it suffices to prove that $p(0; n, k, w) \geq p(0; n, k_0, w)$ when $k < k_0$.

To this end we recall a property of downsets in regular ranked posets. A ranked poset P is *regular* provided that every element of rank k is comparable to the same number of elements of rank $k+1$, and likewise $k-1$. (This is the requirement that each of the bipartite graphs formed by restricting the covering relation of P to two adjacent ranks be regular in the graph theoretic sense.) For example, both the subsets of a set and the subspaces of a vector space, ordered by inclusion, are regular. A *downset* in a partially ordered set is a set S such that $x \in S$ and $y < x$ imply $y \in S$. We claim that if S is a downset then the fraction $|S \cap P_k|/|P_k|$ decreases with k, where P_k is the set of elements of rank k. To see this, let α and β be the common degrees of elements in the bipartite graph $P_k \times P_{k+1}$. Clearly

$$|P_k|\alpha \ = \ \beta|P_{k+1}|. \tag{4}$$

Since S is a downset, every element of $S \cap P_{k+1}$ is related to β elements of $S \cap P_k$. Hence

$$|S \cap P_k|\alpha \ \geq \ \beta|S \cap P_{k+1}|.$$

Dividing the left side by the left side of (4) and the right side by the right side of (4) proves the claim.

Let $\mathcal{I}_j$ be the set of subspaces of $\mathrm{GF}(q)^n$ that contain at most $j(q-1)$ vectors of weight w and no nonzero vectors of less weight. Since this is a downset in the poset of subspaces of $\mathrm{GF}(q)^n$, the fraction of k-dimensional subspaces that lie in $\mathcal{I}_j$ is a decreasing function of k and hence $p(0; n, k, w) \geq p(0; n, k_0, w)$ when $k < k_0$.

$$\text{From now on, we assume that } k \geq k_0 = \lceil n/2 \rceil. \tag{5}$$

(b) Some easy estimates. By (1)

$$\lambda^2 w^2/n < J^2 w^2/n < \mu(n)^2 \to 0 \tag{6}$$

and

$$q^{-d}(ej)^w \;=\; \lambda(q-1)\left(\frac{ejw}{(q-1)n}\right)^w \;\le\; \lambda(q-1)\left(\frac{ejw}{(q-1)n}\right) \;=\; O(J^2 w/n) \tag{7}$$

when $j \le J$. Taking logarithms in (2) and using Stirling's formula, we have

$$\begin{aligned} d - Jw &= w\bigl\{\log_q n - log_q w - J + O(1)\bigr\} - \log_q \lambda \\ &= w\bigl\{(1-B-\epsilon)\log_q n - log_q w + O(1)\bigr\} - O(\log\log n). \end{aligned}$$

Fix K. Since $1 - B - \epsilon \;=\; (1+b-B)/2 \;>\; b$, it follows easily that

$$w\bigl\{(1-B-\epsilon)\log_q n - \log_q w + K\bigr\}$$

is an increasing function of w for sufficiently large n. Hence, for sufficiently large n,

$$d - j(w-1) \ge b\log_q n \quad \text{when} \quad j \le J. \tag{8}$$

(c) Conversion to $d \times n$ Matrices. For V a subspace of $\mathrm{GF}(q)^n$ let $V^\perp$ be its orthogonal complement, the set of vectors orthogonal to every element of V. As noted in the introduction, for nonzero characteristic the intersection $V \cap V^\perp$ may have positive dimension; nevertheless, it is easily checked that the map $V \mapsto V^\perp$ is a bijection from k-dimensional to d-dimensional subspaces. So it suffices to work with $V^\perp$.

Let H be a $d \times n$ matrix whose rows form a basis for $V^\perp$ (in coding theory terminology, a checksum matrix for V). Denote the columns of H by $\mathbf{h}_i$. Note that $\sum v_i \mathbf{h}_i = \mathbf{0}$ if and only if $\mathbf{v} \in V$. If a set S of vectors is linearly dependent and no proper subset is, then call the set minimally dependent. If w is the minimal weight in V, the previous discussion shows that there is a bijection between sets of w minimally dependent vectors among the columns of H and $(q-1)$-tuples of vectors of weight w in V, where a tuple consists of all nonzero multiples of a weight w vector.

Since every d-dimensional subspace of $\mathrm{GF}(q)^n$ has the same number of ordered bases, the fraction of ordered bases with a desired property will be the same as the fraction of d-dimensional subspaces with the property. We will look at ordered bases.

The rows of H are required to be independent; however, the fraction of all $d \times n$ matrices with this property is

$$\begin{aligned} q^{-nd}\prod_{i=0}^{d-1}(q^n - q^i) \;&=\; \prod_{t=k+1}^{n}(1-q^{-t}) \;=\; \exp\Bigl(-\sum_{t=k+1}^{n} q^{-t} + O(q^{-2t})\Bigr) \\ &=\; \exp\bigl(O(q^{-k})\bigr) \;=\; 1 + O(q^{-n/2}), \end{aligned} \tag{9}$$

by (5). We can, in effect, ignore the requirement that the rows of H be independent.

(d) An estimate for $j \le J$. In this part of the proof, we will show that

$$p(j;n,k,w) \;\ge\; (\lambda^j e^{-\lambda}/j!)\{1+O(\mu(n)^2)\} + O(q^{-n/2}) \;\text{ when }\; j \le J. \tag{10}$$

It is instructive, and useful, to treat part of the $w=1$ case separately. A 1-set of minimally dependent vectors must be $\{\mathbf{0}\}$. Hence the fraction of H containing exactly j such sets is

$$q^{-nd}\binom{n}{j}(q^d-1)^{n-j} \;=\; \frac{n^j}{q^{jd}j!}\exp\Big(O(j^2/n)-(n-j)q^{-d}(1+O(q^{-d})\Big)$$

uniformly since $j=o(n^{1/2})$. Using, (9), the theorem now follows easily.

For $w>1$, we generalize this argument. There is a complication: It is now possible for w-sets of minimally dependent vectors to overlap. We count some of the subspaces with nonoverlapping cases and show that this consists of almost all subspaces. Here is how we choose the columns:

(a) Repeat j times: Choose w columns of minimal dependency that are independent of the columns already chosen.

(b) Choose the remaining $n-jw$ columns to avoid introducing more dependent sets of size w or less.

(c) Choose how to order the vectors.

Let N_a and N_c be the number of ways to carry out the choices in (a) and (c). The number of ways to carry out (b) depends on the choices in (a). Let N_b be a lower bound on the number of ways to carry out (b). We seek a lower bound for $N_aN_bN_c$.

Suppose that i sets have been chosen in (a). Since the vectors already chosen span a space of dimension $i(w-1)$, the next set can be chosen in

$$(q^d-q^{i(w-1)})(q^d-q^{i(w-1)+1})\cdots(q^d-q^{i(w-1)+w-2})\times(q-1)^{w-1}$$

ways and so

$$\begin{aligned} N_a \;&=\; \big(q^d(q-1)\big)^{j(w-1)}\prod_{t=0}^{j(w-1)-1}(1-q^{t-d}) \\ &=\; \big(q^d(q-1)\big)^{j(w-1)}\exp\big(O(q^{j(w-1)-d})\big) \;=\; \big(q^d(q-1)\big)^{j(w-1)}\big(1+O(n^{-b})\big) \end{aligned}$$

by (8).

An upper bound on the number of vectors that can be expressed as a linear combination of at most $w-1$ of the vectors in an i-set is $h(i,w)=\sum_{l\le w-1}(q-1)^l\binom{i}{l}$. Thus

$$N_b \;\ge \prod_{i=jw}^{n-1}\Big(q^d-h(i,w)\Big) \;=\; q^{d(n-jw)}\prod_{i=jw}^{n-1}\exp\Big(-q^{-d}h(i,w)+O\big(q^{-2d}h(i,w)^2\big)\Big)$$

provided the expression inside the $O(\ )$ is bounded. Now

$$q^{-d}h(i,w) \ < \ q^{-d}(q-1)^{w-1}\binom{n}{w-1}O(1) \ = \ O(w\lambda/n),$$

and so

$$N_b \ \geq \ q^{d(n-jw)}\exp\Big(-q^{-d}\sum_{l\leq w-1}(q-1)^l\sum_{i=jw}^{n-1}\binom{i}{l} + O(w^2\lambda^2/n)\Big) \tag{11}$$

by (6). Note that for $w \geq 1$, using Stirling's formula,

$$\sum_{l\leq w-1}\binom{jw}{l+1} \ = \ O(1)\binom{jw}{w} \ = \ O((ej)^w).$$

Since

$$\begin{aligned}\sum_{l\leq w-1}(q-1)^l\sum_{i=jw}^{n-1}\binom{i}{l} \ &= \ \sum_{l\leq w-1}(q-1)^l\left(\binom{n}{l+1}-\binom{jw}{l+1}\right)\\ &= \ (q-1)^{w-1}\binom{n}{w}\{1+O(w/n)\}+O((ejq)^w)\\ &= \ \frac{(q-1)^{w-1}n^w}{w!}\{1+O(w^2/n)\}+O((ejq)^w),\end{aligned}$$

we find, using (11), (7), and the definition of λ,

$$\begin{aligned}N_b \ &\geq \ q^{d(n-jw)}\exp\big(-\lambda(1+O(w^2/n))+O(J^2w/n)+O(w^2\lambda^2/n)\big)\\ &= \ q^{d(n-jw)}e^{-\lambda}\big(1+O(J^2w^2/n)\big).\end{aligned}$$

Finally, the number of ways to arrange the vectors, taking into account the fact that there is already some ordering among them is

$$N_c \ = \ \frac{n!}{j!(w!)^j(n-wj)!} \ = \ \frac{n^{wj}}{j!(w!)^j}\big(1+O(w^2j^2/n)\big).$$

Putting all these results together, we obtain the lower bound

$$N_aN_bN_c \ \geq \ q^{nd}\big(\lambda^je^{-\lambda}/j!\big)\big\{1+O(J^2w^2/n)+O(n^{-b})\big\}. \tag{12}$$

Equation (10) follows from (9), (12), and (3).

(e) Completion of the Proof. Summing (10) over $j \leq J$ gives

$$\begin{aligned}1 \ &\geq \ \sum_{j\leq J}p(j;n,k,w) \ \geq \ (1+O(\mu(n)^2))e^{-\lambda}\sum_{j\leq J}\frac{\lambda^j}{j!}+O(Jq^{-n/2})\\ &= \ 1+O(\mu(n)^2)-\sum_{j>J}\frac{\lambda^j}{e^\lambda j!},\end{aligned} \tag{13}$$

where the last equality follows from (3), (5), and the definition of J. Since the ratio of consecutive terms in the last summation is less than

$$\frac{\lambda^{J+1}}{(J+1)!}\Big/\frac{\lambda^J}{J!} \;<\; \lambda/J \;\le\; \frac{B}{B+\epsilon} \;=\; 1-\frac{\epsilon}{B+\epsilon},$$

we have

$$\sum_{j>J}\frac{\lambda^j}{e^\lambda j!} \;<\; \frac{B+\epsilon}{\epsilon}\,\frac{\lambda^J}{e^\lambda J!} \;<\; \frac{B+\epsilon}{\epsilon}\left(\frac{\lambda}{J}\right)^J e^{J-\lambda}.$$

Since this is an increasing function of λ, it is bounded above by $O(D^{\log_q n})$ where

$$D \;=\; \left(\frac{B}{B+\epsilon}\right)^{B+\epsilon} e^\epsilon.$$

For a fixed ϵ, $\left(\frac{B}{B+\epsilon}\right)^{B+\epsilon}$ increases with B; letting B equal 1, and using $0<\epsilon<1/2$ and $1+\epsilon \ge e^{\epsilon-\epsilon^2/2}$, we see

$$D \;\le\; (1+\epsilon)^{-1-\epsilon}e^\epsilon \;\le\; e^{-\epsilon^2/2+\epsilon^3/2} \;\le\; e^{-\epsilon^2/4}.$$

Thus (13) becomes

$$1 \;\ge\; \sum_{j\le J} p(j;n,k,w) \;\ge\; 1+O\big(\mu(n)^2\big)+O(D^{\log_q n}) \;=\; 1+O\big(\mu(n)^2\big), \qquad (14)$$

where the last equality follows from (3). It follows that

$$p(j;n,k,w) \;=\; \frac{\lambda^j e^{-\lambda}}{j!}+O(\mu(n)^2) \;\text{ for }\; j\le J \qquad (15)$$

and $p(j;n,k,w)=O(\mu(n)^2)$ for $j>J$. This completes the proof of Theorem 1.

We now turn to Theorem 2. The standard approximation of the Poisson distribution by the normal says

$$\sqrt{2\pi\lambda}\left(\frac{\lambda^j e^{-\lambda}}{j!}\right) \;-\; e^{-(j-\lambda)^2/2\lambda} \;\to\; 0 \;\text{ uniformly in } j \;\text{ as } \lambda\to\infty.$$

(This can be obtained directly from Stirling's formula.) Since $\lambda^{1/2}\mu(n)^2 \to 0$, Theorem 2 follows from (15).

References

1. M. Aigner, Whitney numbers. In N. White (ed.), *Combinatorial Geometries, volume II*, Cambridge University Press (1987) 139–160.
2. J. R. Goldman and G.-C. Rota, On the foundations of combinatorics, IV: Finite vector spaces and Eulerian generating functions, *Studies Appl. Math.* **49** (1970) 239–258.
3. V. I. Masol, The asymptotics of the number of k-dimensional subspaces of minimal weight over a finite field, *Random Oper. and Stoch. Eqs.* **1** (1993) 287–292.