STEIN’S METHOD AND THE RANK DISTRIBUTION OF RANDOM MATRICES OVER FINITE FIELDS

JASON FULMAN AND LARRY GOLDSTEIN

Abstract. With $Q_{q,n}$ the distribution of $n$ minus the rank of a matrix $U_n$ chosen uniformly from $\text{Mat}(n,q)$, the collection of all $n \times n$ matrices over the finite field $\mathbb{F}_q$ of size $q \geq 2$, and $Q_q$ the distributional limit of $Q_{q,n}$ as $n \to \infty$, we apply Stein’s method to prove the total variation bound

$$\frac{1}{8q^{n+1}} \leq ||Q_{q,n} - Q_q||_{TV} \leq \frac{3}{q^{n+1}}.$$

1. Introduction

We study the distribution of the rank of a matrix $U_n$ chosen uniformly from $\text{Mat}(n,q)$, the collection of all $n \times n$ matrices over the finite field $\mathbb{F}_q$ of size $q \geq 2$. It is known (page 38 of [3]) that for all $k$ in the integer interval $\mathbb{I}_n = \{0, \ldots, n\}$,

(1) $P(\text{rank}(U_n) = n - k) = p_{k,n}$ where

$$p_{k,n} = \frac{\prod_{j=k+1}^{n}(1-q^{-j})^2}{q^k \prod_{j=1}^{n-k}(1-q^{-j})}.$$

Clearly for any fixed $k \in \mathbb{N}_0$,

(2) $\lim_{n \to \infty} p_{k,n} = p_k$ where

$$p_k = \frac{\prod_{i \geq 1}(1-q^{-i})}{q^k \prod_{j=1}^{k}(1-q^{-j})^2}.$$

Our main result, Theorem 1.1, provides sharp upper and lower bounds on the total variation distance between the distribution in (1) of $n$ minus the rank of $U_n$, denoted $Q_{q,n}$, and its limit in (2), denoted $Q_q$. Recall that the total variation distance between two probability distributions $P_1, P_2$ on a finite set $S$ is given by

(3) $||P_1 - P_2||_{TV} := \frac{1}{2} \sum_{s \in S} |P_1(s) - P_2(s)| = \max_{A \subseteq S} |P_1(A) - P_2(A)|.$

Theorem 1.1. For $q \geq 2$ and $n \geq 1$,

(4) $\frac{1}{8q^{n+1}} \leq ||Q_{q,n} - Q_q||_{TV} \leq \frac{3}{q^{n+1}}.$
The upper bound in Theorem 1.1 appears quite difficult to compute directly by substituting the expressions for the point probabilities given in (1) and (2) into the defining expressions for the total variation distance in (3). On the other hand, use of Stein’s method [19], [7] makes for a quite tractable computation. The next paragraph gives pointers to the large literature on ranks of random matrices over finite fields. Our hope is that the Stein’s method machinery developed here for bounding $||Q_{q,n} - Q_q||_{TV}$ may prove useful for other problems about ranks of random matrices over finite fields.

First, perhaps the earliest systematic study of ranks of random matrices from the finite classical groups is due to Rudvalis and Shinoda [17], [18]. They determine the rank distribution of random matrices from finite classical groups, and relate distributions such as $Q_q$ to identities of Euler. Second, ranks of random matrices from finite classical groups appear in works on the “Cohen-Lenstra heuristics” of number theory; see [22] for the finite general linear groups and [15] for the finite symplectic groups. Third, ranks of random matrices over finite fields are useful in coding theory; see [4] and Chapter 15 of [14]. Fourth, there is work on ranks of random matrices over finite fields where the matrix entries are independent and identically distributed, but not necessarily uniform. For example the paper [6] uses a combination of Möbius inversion, finite Fourier transforms, and Poisson summation, to find conditions on the distribution of matrix entries under which the probability of a matrix being invertible tends to $p_0$ as $n \to \infty$. Further results in this direction, including rank distributions of sparse matrices, can be found in [5], [8], [9].

The organization of this paper is as follows. Section 2 shows that if $X$ is from the distribution $Q_{q,n}$, then $E(q^X) = 2 - 1/q^n$. This seems challenging to prove directly from the definition of $Q_{q,n}$; in fact even proving that $Q_{q,n}$ is a probability distribution is tricky. To overcome this obstacle, we use the interpretation of $Q_{q,n}$ in terms of ranks of random matrices, together with some combinatorial arguments. Section 3 sets up the necessary machinery to apply Stein’s method, in particular, finding a characterizing equation for the target distribution which yields a “Stein” equation, and then bounding the norm of its solutions. The path followed here is along the lines of the “comparison of generators” approach as in [12] and [13]. The Stein apparatus developed in Section 3 is then applied with the result from Section 2 to yield Theorem 1.1.

2. A PROPERTY OF $Q_{q,n}$

The purpose of this section is to prove Theorem 2.1. The proof assumes familiarity with rational canonical forms of matrices (that is the theory of Jordan forms over finite fields), and with cycle index generating functions. Background on these topics can be found in [10] or [20], or in the survey [11].
Theorem 2.1. If $X$ has the $Q_{q,n}$ distribution on $\{0, 1, \ldots, n\}$, then

$$E(q^X) = 2 - 1/q^n.$$ 

Proof. The sought equation is

$$\sum_{k=0}^{n} q^k p_{k,n} = 2 - 1/q^n. \tag{5}$$

From the expression for $p_{k,n}$ in (1), it is clear that if one multiplies (5) by $q^t(1-1/q) \cdots (1-1/q^n)$ where $t$ is sufficiently large as a function of $n$, then both sides become polynomials in $q$. Since polynomials in $q$ agreeing for infinitely many values of $q$ are equal, it is enough to prove the result for infinitely many values of $q$, so we demonstrate it for $q$ a prime power.

Let $M$ be an $n \times n$ matrix over $\mathbb{F}_q$. Then $n$ minus the rank of $M$ is equal to $l(\lambda_z(M))$, the number of parts in the partition corresponding to the $z$-piece of the rational canonical form of $M$. Thus

$$E(q^X) = \frac{1}{q^n} \sum_{M \in \text{Mat}(n,q)} q^{l(\lambda_z(M))}, \tag{6}$$

where $\text{Mat}(n,q)$ denotes the set of $n \times n$ matrices over the finite field $\mathbb{F}_q$.

From the cycle index for $\text{Mat}(n,q)$ ([20]), it follows that

$$1 + \sum_{n \geq 1} \frac{u^n}{|GL(n,q)|} \sum_{M \in \text{Mat}(n,q)} q^{l(\lambda_z(M))} = \left[ \sum_{\lambda} q^{l(\lambda)} u^{\lambda} \right] \prod_{\phi \neq z} \sum_{\lambda} u^{\lambda|\deg(\phi)} c_{GL,\phi}(\lambda). \tag{7}$$

Here $\lambda$ ranges over all partitions of all natural numbers, and $l(\lambda)$ is the number of parts of $\lambda$. The quantity $c_{GL,\phi}(\lambda)$ is a certain function of $\lambda, \phi$ which depends on $\phi$ only through its degree. The product is over all monic, irreducible polynomials $\phi$ over $\mathbb{F}_q$ other than $\phi = z$.

From the cycle index for $GL(n,q)$ ([20]), it follows that

$$\frac{1}{1-u} = 1 + \sum_{n \geq 1} \frac{u^n}{|GL(n,q)|} \sum_{\alpha \in GL(n,q)} 1 = \prod_{\phi \neq z} \sum_{\lambda} u^{\lambda|\deg(\phi)} c_{GL,\phi}(\lambda). \tag{8}$$

Summarizing, it follows from (7) and (8) that

$$1 + \sum_{n \geq 1} \frac{u^n}{|GL(n,q)|} \sum_{M \in \text{Mat}(n,q)} q^{l(\lambda_z(M))} = \frac{1}{1-u} \sum_{\lambda} q^{l(\lambda)} u^{\lambda|\deg(\phi)} \sum_{\phi \neq z} c_{GL,\phi}(\lambda). \tag{9}$$

The next step is to compute

$$\sum_{\lambda} q^{l(\lambda)} u^{\lambda|\deg(\phi)} c_{GL,\phi}(\lambda) = \sum_{\lambda} q^{l(\lambda)} u^{\lambda|\deg(\phi)} c_{GL,\phi-1}(\lambda).$$
From the cycle index of $GL(n, q)$, it follows that

$$1 + \sum_{n=1}^{\infty} \frac{u^n}{|GL(n, q)|} \sum_{\alpha \in GL(n, q)} q^{l(\lambda_{z-1}(\alpha))} = \sum_{\lambda} q^{l(\lambda)} u^{\lambda} \prod_{\phi \neq z} \frac{u^{\lambda(\phi)}}{c_{GL, \phi}(\lambda)}$$

$$= \sum_{\lambda} \frac{q^{l(\lambda)} u^{\lambda}}{c_{GL, z-1}(\lambda)} \prod_{\phi \neq z} \frac{u^{\lambda(\phi)}}{c_{GL, \phi}(\lambda)}$$

$$= \frac{1}{1 - u} \sum_{\lambda} \frac{q^{l(\lambda)} u^{\lambda}}{c_{GL, z-1}(\lambda)}$$

$$= \frac{\prod_{i \geq 1} (1 - u/q^i)}{1 - u} \sum_{\lambda} \frac{q^{l(\lambda)} u^{\lambda}}{c_{GL, z-1}(\lambda)}.$$ 

The third equality used (8) and the final equality is from Lemma 6 of [20] and page 19 of [1].

Next we can use group theory to compute

$$1 + \sum_{n=1}^{\infty} \frac{u^n}{|GL(n, q)|} \sum_{\alpha \in GL(n, q)} q^{l(\lambda_{z-1}(\alpha))} = 1 + \sum_{n \geq 1} 2u^n = \frac{1 + u}{1 - u}.$$ 

Indeed, $q^{l(\lambda_{z-1}(\alpha))}$ is the number of fixed points of $\alpha$ in its action on the underlying $n$ dimensional vector space $V$. By Burnside’s lemma (page 95 of [21]), the average number of fixed points of a finite group acting on a finite set is the number of orbits of the action on the set. For $GL(n, q)$ acting on $V$, there are two such orbits, consisting of the zero vector and the set of non-zero vectors. Thus

(10) $1 + \sum_{n=1}^{\infty} \frac{u^n}{|GL(n, q)|} \sum_{\alpha \in GL(n, q)} q^{l(\lambda_{z-1}(\alpha))} = 1 + \sum_{n \geq 1} 2u^n = \frac{1 + u}{1 - u}.$

By the previous two paragraphs,

(11) $\sum_{\lambda} q^{l(\lambda)} u^{\lambda} = \frac{1 + u}{\prod_{i \geq 1} (1 - u/q^i)}$

It follows from (9) and (11) that

$$1 + \sum_{n \geq 1} \frac{u^n}{|GL(n, q)|} \sum_{M \in Mat(n, q)} q^{l(\lambda_{z}(M))} = \frac{1 + u}{1 - u} \prod_{i \geq 1} \frac{1}{1 - u/q^i}.$$ 

Thus by (6), $E(q^X)$ is $\frac{|GL(n, q)|}{q^n}$ multiplied by the coefficient of $u^n$ in

$$\frac{1 + u}{1 - u} \prod_{i \geq 1} \frac{1}{1 - u/q^i}.$$
From page 19 of [1], the coefficient of $u^n$ in
\[
\frac{1}{1-u} \prod_{i \geq 1} \frac{1}{1-u/q^i}
\]
is equal to \(\frac{1}{(1-1/q)(1-1/q^2) \cdots (1-1/q^n)}\). Thus,
\[
E(q^X) = \left| GL(n,q) \right| q^{n^2} \left[ \frac{1}{(1-1/q) \cdots (1-1/q^n)} + \frac{1}{(1-1/q) \cdots (1-1/q^{n-1})} \right]
\]
\[= 2 - \frac{1}{q^n},\]
where the last equality used that \(\left| GL(n,q) \right| = q^{n^2} (1-1/q) \cdots (1-1/q^n)\). □

We close this section with two remarks about the distribution $Q_{q,n}$.

(1) From [3], there is a natural Markov chain on \(\{0, 1, \cdots, n\}\) which has $Q_{q,n}$ as its stationary distribution. This chain has transition probabilities
\[
M(i, i+1) = \frac{q^{n-i-1}(q^n-1)}{(q^n-1)^2}, \quad M(i, i-1) = \frac{(q^n-q^{n-i})^2}{(q^n-1)^2}
\]
\[
M(i, i) = 1 - M(i, i-1) - M(i, i+1)
\]
This Markov chain describes how the rank of a matrix evolves by adding a uniformly chosen rank one matrix at each step.

(2) It is known (page 338 of [21]) that the number of $n \times n$ matrices over the finite field $\mathbb{F}_q$ with rank $r$ is equal to
\[
\left[ \begin{array}{c} n \\ r \end{array} \right] q \sum_{k=0}^{r} (-1)^{r-k} \left[ \begin{array}{c} r \\ k \end{array} \right] q^{nk+(r-k)}.
\]
Here
\[
\left[ \begin{array}{c} n \\ m \end{array} \right]_q = \frac{(q^n-1)(q^{n-1}-1) \cdots (q^{n-m+1}-1)}{(q^{m-1}-1)(q^{m-2}-1) \cdots (q-1)}
\]
is the $q$-binomial coefficient.

Since the proportion of $n \times n$ matrices over $\mathbb{F}_q$ with rank $r$ is also given by $p_{n-r,n}$, we obtain the following corollary, which we have not seen stated in the literature.

**Corollary 2.2.** For $0 \leq r \leq n$,
\[
\frac{\prod_{j=n-r+1}^{n}(1-q^{-j})^2}{q^{(n-r)^2} \prod_{j=1}^{r}(1-q^{-j})} = \frac{1}{q^{n^2}} \left[ \begin{array}{c} n \\ r \end{array} \right] q \sum_{k=0}^{r} (-1)^{r-k} \left[ \begin{array}{c} r \\ k \end{array} \right] q^{nk+(r-k)}.
\]
3. Stein’s method computations

We begin with a general result for obtaining characterizations of discrete integer distributions. Our resulting identities are in the spirit of Proposition 2.1 and Corollary 2.1 of [12], but of a somewhat simpler form not involving forward differences. We say a nonempty subset $\mathbb{I}$ of the integers $\mathbb{Z}$ is an interval if $a, b \in \mathbb{I}$ with $a \leq b$ then $[a, b] \cap \mathbb{Z} \subseteq \mathbb{I}$. Let $\mathcal{L}(X)$ denote the distribution of a random variable $X$.

**Lemma 3.1.** Let $\{r_k, k \in \mathbb{I}\}$ be the distribution of a random variable $Y$ having support the integer interval $\mathbb{I}$. Then for any functions $a(k)$ and $b(k)$ satisfying

$$\frac{b(k)}{a(k)} = \frac{r_{k-1}}{r_k} \text{ for all } k \in \mathbb{I},$$

a random variable $X$ with support $\mathbb{I}$ has distribution $\mathcal{L}(Y)$ if and only if

$$E[a(X+1)f(X+1)] = E[b(X)f(X)]$$

(12)

for all functions $f$ for which the expectations in (12) exist.

**Proof.** If $\mathcal{L}(X) = \mathcal{L}(Y)$ then for all $k \in \mathbb{I}$

$$\frac{P(X = k-1)}{P(X = k)} = \frac{r_{k-1}}{r_k} = \frac{b(k)}{a(k)},$$

so that

$$E(a(X+1)1(X + 1 = k)) = a(k)P(X = k - 1) = b(k)P(X = k) = E(b(X)1(X = k)).$$

Hence (12) holds for $f(x) = 1(x = k)$. By linearity, (12) holds for all functions with finite support, and hence for all the claimed functions by dominated convergence.

Conversely, if (12) holds for $X$ then setting $f(x) = 1(x = k)$ for $k \in \mathbb{I}$ yields

$$a(k)P(X = k - 1) = b(k)P(X = k),$$

so that

$$\frac{P(X = k-1)}{P(X = k)} = \frac{b(k)}{a(k)} = \frac{r_{k-1}}{r_k}.$$ 

As $X$ has the same integer interval support $\mathbb{I}$ as $Y$, we conclude that $\mathcal{L}(X) = \mathcal{L}(Y)$. \qed

For example, when $Y \sim \mathcal{P}(\lambda)$, the Poisson distribution with parameter $\lambda$, we have

$$\frac{r_{k-1}}{r_k} = \frac{k}{\lambda}.$$
and setting \( b(k) = k \) and \( a(k) = \lambda \) yields the standard characterization of the Poisson distribution [2],

\[
E[\lambda f(Y + 1)] = E[Y f(Y)].
\]

**Lemma 3.2.** If \( Q \) has the \( Q_q \) distribution then

\[
E[q f(Q + 1)] = E[(q^Q - 1)^2 f(Q)]
\]

for all functions \( f \) for which these expectations exist.

If \( Q_n \) has the \( Q_{q,n} \) distribution then

\[
E[q(1 - q^{-n+Q_n})f(Q_n + 1)] = E[(q^{Q_n} - 1)^2 f(Q_n)]
\]

for all functions \( f \) for which these expectations exist.

**Proof.** From (2) we obtain

\[
\frac{p_{k-1}}{p_k} = \frac{(q^k - 1)^2}{q} \quad \text{for all } k \in \mathbb{N}_0.
\]

Applying Lemma 3.1 with \( a(k) = q \) and \( b(k) = (q^k - 1)^2 \) yields (13). Similarly, from (1) we obtain

\[
\frac{p_{k-1,n}}{p_{k,n}} = \frac{(q^k - 1)^2}{q(1 - q^{-n+k-1})} \quad \text{for all } k \in \mathbb{I}_n,
\]

and we set \( a(k) = q(1 - q^{-n+k-1}) \) and \( b(k) = (q^k - 1)^2 \) in Lemma 3.1 to achieve (14).

For any given \( Y \) with support \( \mathbb{I} \), one may always apply Lemma 3.1 with \( a(k) = \mu \), some nonzero constant, and \( b(k) = \mu r_{k-1}/r_k \) for \( k \in \mathbb{I} \), resulting in the characterization

\[
\mu E[f(Y + 1)] = \mu E\left[\frac{r_Y-1}{r_Y} f(Y)\right],
\]

and the associated Stein equation

\[
\mu \left(f(k + 1) - \frac{r_{k-1}}{r_k} f(k)\right) = h(k) - \mathcal{Y} h
\]

for functions \( h \) such that the \( \mathcal{Y} h \) expectation, that is \( E h(Y) \), of \( h \) exists.

For the special case of \( Q \) with distribution \( Q_q \) given in (2), the Stein equation (16) with \( \mu = q \) becomes

\[
q f(k + 1) - (q^k - 1)^2 f(k) = h(k) - Q_q h \quad k \in \mathbb{N}_0.
\]

As the coefficient of \( f(0) \) in (17) is zero when \( k = 0 \), we may arbitrarily set \( f(0) = 0 \). It is easy to verify that for every \( h : \mathbb{N}_0 \to \mathbb{R} \) such that \( Q h \) exists the solution to (17) is given by

\[
f(k+1) = \frac{1}{qp_k} \sum_{j=0}^{k} p_j [h(j) - Q h] = \frac{E[(h(Q) - Q h)1(Q \leq k)]}{qp_k} \quad \text{for } k \in \mathbb{N}_0.
\]
In particular, for \( h_A(k) = 1(k \in A) \) with \( A \subset \mathbb{N}_0 \) and \( U_k = \{0, 1, \ldots, k\} \), following the argument of Barbour et al [2], Lemma 1.1.1, for \( k \in \mathbb{N}_0 \) we may write the solution \( f_A \) as

\[ f_A(k+1) = \frac{P(Q \in A \cap U_k) - P(Q \in A)P(Q \in U_k)}{qp_k} \]

\[ = \frac{P(Q \in A \cap U_k)P(Q \in U_k^c) - P(Q \in A \cap U_k^c)P(Q \in U_k)}{qp_k}. \tag{19} \]

For a function \( f : \mathbb{N}_0 \to \mathbb{R} \), let

\[ ||f|| = \sup_{k \in \mathbb{N}_0} |f(k)|. \]

The following lemma is crucial, and should be useful for other applications of Stein’s method to the distribution \( Q_q \).

**Lemma 3.3.** For \( q \geq 2 \) and \( A \subset \mathbb{N}_0 \), if \( f_A \) is the solution (19) to the Stein equation (17) for \( h_A(k) = 1(k \in A) \), then

\[ \sup_{A \subset \mathbb{N}_0} ||f_A|| \leq \frac{1}{q^2} + \frac{1}{q^3}. \]

**Proof.** As we may set \( f_A(0) = 0 \) it suffices to consider \( f_A(k+1) \) for \( k \in \mathbb{N}_0 \). By (19), we obtain

\[ |f_A(k+1)| \leq \frac{P(Q \in U_k)P(Q \in U_k^c)}{qp_k} \tag{20} \]

with equality when \( A = U_k \).
By neglecting the term \(P(Q \in U_k)\) in (20) we obtain

\[
|f_A(k+1)| \leq \frac{P(Q \in U_k^c)}{qp_k} = q^{k^2-1} \prod_{j=1}^{k} \left(1 - \frac{1}{q^j}\right)^2 \sum_{l=k+1}^{\infty} \frac{1}{q^l} \prod_{j=k+1}^{l} \left(1 - \frac{1}{q^j}\right)^2
\]

\[
= q^{k^2-1} \sum_{l=k+1}^{\infty} \frac{1}{q^2} \prod_{j=k+1}^{l} \left(1 - \frac{1}{q^j}\right)^2
\]

\[
\leq \prod_{j=k+1}^{\infty} \left(1 - \frac{1}{q^j}\right)^2 \sum_{l=k+1}^{\infty} \frac{1}{q^l} 
\]

\[
= \frac{q^{k^2-1}}{(1 - \sum_{j=k+1}^{\infty} \frac{1}{q^j})^2} \sum_{l=k+1}^{\infty} \frac{1}{q^l} 
\]

\[
= \frac{q^{k^2}}{q(1 - \frac{1}{q^k})^2} \sum_{l=1}^{\infty} \frac{1}{q^{l(k+1)^2}} 
\]

\[
= \frac{1}{q(1 - \frac{1}{q^k(q-1)})^2} \sum_{l=1}^{\infty} \frac{1}{q^{2l^2}} 
\]

\[
\leq \frac{1}{q(q^k - \frac{1}{q-1})^2} \sum_{l=1}^{\infty} \frac{1}{q^{l^2}},
\]

where for the second inequality we are using that fact, easily shown by induction, that \(\prod_{i=1}^{n} (1 - a_i) \geq 1 - \sum_{i=1}^{n} a_i\) whenever \(a_i \in [0, 1], i = 1, \ldots, n\).

For \(k \geq 1\), using \(q \geq 2\), we obtain

\[
|f_A(k+1)| \leq \frac{1}{q(q - \frac{1}{q-1})^2} \sum_{l=1}^{\infty} \frac{1}{q^{l^2}} \leq \frac{4}{q^3} \sum_{l=1}^{\infty} \frac{1}{q^{l^2}}.
\]

From (19) with \(k = 0\) we obtain

\[
f_A(1) = \frac{P(Q \in A \cap U_0)P(Q \geq 1) - P(Q \in A \cap U_0^c)P(Q = 0)}{qp_0}.
\]

If \(A \ni 0\) then

\[
|f_A(1)| = \left|\frac{P(Q = 0)P(Q \geq 1) - P(Q \in A \setminus \{0\})P(Q = 0)}{qp_0}\right| = \left|\frac{P(Q \geq 1) - P(Q \in A \setminus \{0\})}{q}\right| \leq \frac{P(Q \geq 1)}{q},
\]

while if \(A \notni 0\) then again

\[
|f_A(1)| = \frac{P(Q \in A)P(Q = 0)}{qp_0} \leq \frac{P(Q \geq 1)}{q},
\]

\[
\]
so that for all $A \subset \mathbb{N}_0$

$$|f_A(1)| \leq \frac{P(Q \geq 1)}{q} = \frac{1 - p_0}{q}$$

$$= \frac{1}{q}\left(1 - \prod_{i \geq 1}(1 - \frac{1}{q^i})\right)$$

$$\leq \frac{1}{q}\left(\frac{1}{q} + \frac{1}{q^2}\right)$$

$$= \frac{1}{q^2} + \frac{1}{q^3},$$

where we have applied the inequality $\prod_{i \geq 1}(1 - \frac{1}{q^i}) \geq 1 - 1/q - 1/q^2$, valid for all $q \geq 2$, obtained by taking limits in Lemma 3.5 of [16]. Again using $q \geq 2$ we obtain

$$\sup_{k \in \mathbb{N}_0} |f_A(k + 1)| \leq \max \left\{\frac{4}{q^3} \sum_{l=1}^{\infty} \frac{1}{q^l}, \frac{1}{q^2} + \frac{1}{q^3}\right\} = \frac{1}{q^2} + \frac{1}{q^3},$$

completing the proof. $\square$

Next we present the proof of our main result, Theorem 1.1.

**Proof.** We first compute the lower bound on the total variation distance by estimating the difference of the two distributions at $k = 0$. In particular, by (3), (1) and (2),

$$||Q_{q,n} - Q_q||_{TV} \geq \frac{1}{2} [p_{0,n} - p_0]$$

$$= \frac{1}{2} \left[\prod_{1 \leq i \leq n} (1 - 1/q^i) - \prod_{i \geq 1} (1 - 1/q^i)\right]$$

$$\geq \frac{1}{2} \left[(1 - 1/q) \cdots (1 - 1/q^n) - (1 - 1/q) \cdots (1 - 1/q^{n+1})\right]$$

$$= \frac{1}{2q^{n+1}} (1 - 1/q) \cdots (1 - 1/q^n)$$

$$\geq \frac{1}{8q^{n+1}} (1 - 1/q - 1/q^2)$$

The third inequality used Lemma 3.5 of [16], which states that for $q \geq 2$, $(1 - 1/q) \cdots (1 - 1/q^n) \geq (1 - 1/q - 1/q^2)$.

For the upper bound, let $Q_n$ and $Q$ have the $Q_{q,n}$ and $Q_q$ distributions respectively. For $A \subset \mathbb{N}_0$ let $f_A$ be the solution (19) to the Stein equation
(17) for \( h_A(k) = 1(k \in A) \). Then

\[
|P(Q_n \in A) - P(Q \in A)| = |E[h_A(Q_n)] - Q_q h_A| \\
= |E[qf_A(Q_n + 1) - (q^{Q_n} - 1)^2 f_A(Q_n)]| \\
= |E[q^{-n+Q_n+1} f_A(Q_n + 1)]| \\
\leq ||f_A|| E q^{-n+Q_n+1},
\]

where we have applied (14) in the third equality. Applying Theorem 2.1 and Lemma 3.3 gives that

\[
||f_A|| E q^{-n+Q_n+1} \leq \left( \frac{1}{q^2} + \frac{1}{q^3} \right) q^{-n+1} \left( 2 - \frac{1}{q} \right) \leq \frac{2 \left( 1 + \frac{1}{q} \right)}{q^{n+1}} \leq 3 q^{n+1}.
\]

Now taking the supremum over all \( A \subset \mathbb{N}_0 \) and applying definition (3) completes the proof.

4. Acknowledgements

Fulman was supported by a Simons Foundation Fellowship. Goldstein was supported by NSA grant H98230-11-1-0162.

References


Department of Mathematics, University of Southern California, Los Angeles, CA, 90089

E-mail address: fulman@usc.edu

Department of Mathematics, University of Southern California, Los Angeles, CA, 90089

E-mail address: larry@math.usc.edu