

# On the missing values of $n! \bmod p$

Kevin A. Broughan and A. Ross Barnett

*University of Waikato, Hamilton, New Zealand*

*Version: 15th Jan 2009*

E-mail: kab@waikato.ac.nz, arbus@math.waikato.ac.nz

We show that the average proportion of missing values of a sequence of  $N$  components with  $N$  values, and with no equal neighbors, is  $1/e$ .

*Key Words:* factorial, missing values.

MSC2000: 11A41, 11B05, 11B83.

## 1. INTRODUCTION

Many authors have commented on the question of measuring the missing residue classes of the sequence of factorials taken modulo primes, [1, 2, 3, 4, 5]. When looked at as a function of primes  $p$ , the number of missing classes, denoted  $N_p$ , cluster about a line with slope close to  $1/e$  and scatter of the order of  $\sqrt{p}$ . This clustering is quite striking. See Figure 1 which is a plot of points  $(p, N_p)$  for primes  $p$  with  $2 \leq p \leq 8000$ .

In spite of this clear numerical evidence, analytic evidence is almost entirely absent. That  $N_p \leq p - \sqrt{p-1}$  follows from the relation

$$n = n!((n-1)!)^{-1} \bmod p,$$

whereas Banks, Luca, Shparlinski and Stichenoth have shown in [1] that  $\limsup_{p \rightarrow \infty} N_p = \infty$ . Neither of these results sheds light on why the line and the slope  $1/e$  should appear.

In contrast to this small progress, Cobeli, Vâjâitu and Zaharescu [2] reveal that the slope  $1/e$  appears quite naturally when the set of all sequences with  $N$  values and of length  $N$  is considered with  $N$  sufficiently large. The proportion of missing values is measured “on average” to be  $1/e$ . In some sense, the factorial sequence is “generic”. But this still leaves the fun-

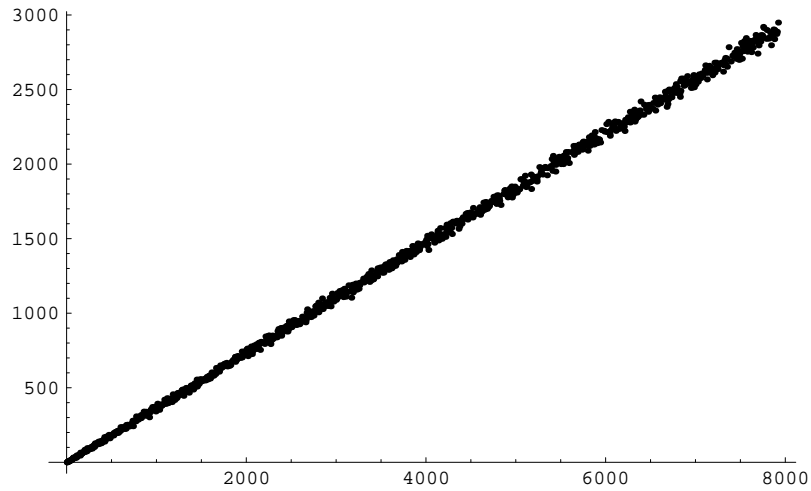


FIG. 1. The number of missing values of  $n! \bmod p$  as a function of  $p$ .

damental question regarding the missing values of the factorial sequence unanswered.

In this paper we take a step closer to the factorial sequence using Cobeli, Vâjăitu and Zaharescu's method. Here we do not consider all sequences but only those which share a property with  $n! \bmod p$ , namely, if we let the sequence of least positive residues of the factorials be  $(x_1, \dots, x_{p-1})$ , so  $i! \equiv x_i \pmod p$ , then because  $(i+1)! \not\equiv i! \pmod p$ , it follows that  $x_{i+1} \neq x_i$  for all  $i$  with  $1 \leq i \leq p-2$ . Sequences with this **no-equal-neighbors** condition number less than half of all sequences (Lemma 2.1). When we average over this subset we also obtain the slope value  $1/e$  (Theorem 2.1, Corollary).

A number of additional constraints could be included, such as  $x_1 = 1, x_2 = 2$ , but these make no essential difference to the discussion or to the results. However, constraints in addition to "no-equal-neighbors" such as "no more than  $n$  components which are equal should be  $n$  apart" with  $n = 2, 3, \dots$  provide difficult counting problems and will have to await further work.

In what follows we use the notation of [2].

## 2. RESULTS

Let  $M$  and  $N$  be given natural numbers. First we derive a formula which implies that, in the case  $M = N$ , less than half of the sequences satisfy the no-equal-neighbors condition.

LEMMA 2.1. *Let  $M$  and  $N$  be natural numbers and let  $S(M, N) \subset M^N$  be the set of sequences  $x = (x(n) : 1 \leq n \leq N)$  with  $x(n) \neq x(n+1)$  for  $1 \leq n \leq N-1$ . (Note that these sequences include the factorials modulo  $N$ .) Then the number of elements in  $S(M, N)$  is given by*

$$|S(M, N)| = M \cdot (M-1)^{N-1}.$$

*Proof.* By inclusion-exclusion:

$$\begin{aligned} |S(M, N)| &= M^N - |\{x : \text{there exists at least one } n \text{ with } x(n) = x(n+1)\}| \\ &\quad + |\{x : \text{there exists at least two } n\text{'s with } x(n) = x(n+1)\}| - \dots \\ &= M^N - \binom{N-1}{1} M^{N-1} + \binom{N-1}{2} M^{N-2} \\ &\quad - \dots + (-1)^{N-1} \binom{N-1}{N-1} M^1 = M \cdot (M-1)^{N-1}. \end{aligned}$$

■

It follows from the lemma that if  $M, N$  tend to infinity in such a way that  $N/M \rightarrow \lambda$  then

$$\lim_{M, N \rightarrow \infty} \frac{|S(M, N)|}{M^N} = \frac{1}{e^\lambda}.$$

Next we see how to count the number of subsets of  $S \subset \{1, \dots, n\}$  with the restriction that if  $i \in S$  then  $i-1 \notin S$  and  $i+1 \notin S$ .

DEFINITION 2.1. For natural numbers  $n$  and  $j$ , let  $L(n, j)$  be the number of subsets of  $\{1, 2, \dots, n\}$  of size  $j \geq 1$  such that no two elements, in the natural ordering, are neighbors. Call this the **no-neighbors** condition. For example  $L(3, 2) = 1$  because  $\{1, 3\}$  is the only admissible subset of  $\{1, 2, 3\}$ .

LEMMA 2.2. *For  $n, j \geq 3$  the function  $L(n, j)$  satisfies the recurrence*

$$L(n, j) = L(n-1, j) + L(n-2, j-1).$$

It has values which may be written

- (1)  $L(n, 1) = n$ ,
- (2)  $L(n, 2) = 0$  for  $1 \leq n \leq 2$ ,
- (3)  $L(n, 2) = \binom{n-1}{2}$  for  $n \geq 3$ ,  
and for  $j \geq 3$ ,
- (4)  $L(n, j) = 0$  for  $1 \leq n \leq 2j - 1$ ,
- (5)  $L(n, j) = \binom{n-j+1}{j}$  for  $n \geq 2j$ .

*Proof.* Items (1) and (2) follow directly. For (3) order the elements of  $\{1, 2, \dots, n\}$  from the left as usual. If the first element chosen is in position  $j$  then there are  $n - j - 1$  ways of choosing the second element. Thus the number of subsets is

$$(n-2) + (n-3) + \dots + 1 = \frac{(n-1)(n-2)}{2}.$$

Item (4) follows because any subset of size  $j$  with  $2j > n$  has at least one set of neighbors. To see (5), divide the subsets up into two disjoint families as follows:

$$\begin{aligned} L(n, j) &= |\{\text{subsets which do not include the leftmost element}\}| \\ &\quad + |\{\text{subsets which do include that element}\}| \\ &= L(n-1, j) + L(n-2, j-1). \end{aligned}$$

Then the expression for  $L(n, j)$  follows from the standard binomial recurrence. ■

Note that if  $u_n$  is the number of non-empty subsets of  $\{1, \dots, n\}$  satisfying the no-neighbors condition then  $u_3 = 4$ ,  $u_4 = 7$  and for  $n > 4$ ,

$$u_n = u_{n-1} + u_{n-2} + \frac{3 + (-1)^n}{2}.$$

This relation is not needed subsequently.

A key technical lemma relating the number of subsets satisfying the no-neighbors condition and the set of all subsets is the following:

LEMMA 2.3. *If  $M, N \rightarrow \infty$  in such a way that  $N/M \rightarrow \lambda \leq 1$  then*

$$\lim_{M, N \rightarrow \infty} \frac{\sum_{j=0}^{N/2} \binom{N-j+1}{j} \left(\frac{-1}{M}\right)^j}{\left(1 - \frac{1}{M}\right)^N} = 1.$$

*Proof.* First expand the binomial coefficient for the absolute value of a single term in the sum in the numerator:

$$\binom{N-j+1}{j} \frac{1}{M^j} = \frac{(\frac{N}{M} - \frac{j-1}{M})(\frac{N}{M} - \frac{j}{M}) \cdots (\frac{N}{M} - \frac{2j-2}{M})}{j!}.$$

Neglecting the negative terms in each factor in the numerator leads to the upper bound  $(N/M)^j/j!$ . Summing one negative term from each times the  $(j-1)$ 'th power of the leading term gives (justified, say, using induction) the lower bound

$$\frac{(N/M)^j}{j!} - \frac{1}{M} \frac{3j(j-1)}{2j!} \left(\frac{N}{M}\right)^{j-1}.$$

Including the alternating signs and adding gives the limit

$$\lim_{M, N \rightarrow \infty} \sum_{j=0}^{N/2} \binom{N-j+1}{j} \left(\frac{-1}{M}\right)^j = e^{-\lambda}$$

from which the stated result follows.  $\blacksquare$

Now we begin to count sequences. First focus on the number of vectors associated with a gap of length  $g$  between two specified fixed values  $y$ .

**DEFINITION 2.2.** For  $g \geq 1$  let  $P_g(M)$  be the number of possible choices for vectors in  $M^N$  satisfying the no-equal-neighbors condition, and having two fixed values  $y \in \mathcal{N}$  separated by a gap of length  $g$ , where the vectors are free to take any values in  $\{1, \dots, M\}$ , including the value  $y$ , provided the no-equal-neighbors condition continues to be satisfied.

**LEMMA 2.4.** For all  $g \geq 1$  let  $Q_g(M) := (M-1) \cdot (M-2)^{g-1}$ . Then (1)  $P_1(M) = (M-1)$ , (2)  $P_2(M) = (M-1)(M-2)$  and, (3) for  $g \geq 3$

$$P_g(M) := Q_1(M)P_{g-2}(M) + Q_2(M)P_{g-3}(M) + \cdots + Q_{g-2}(M)P_1(M) + Q_g(M).$$

Furthermore, for all  $g \geq 1$

$$\begin{aligned} P_g(M) &= \frac{(M-1)^{g+1} + (-1)^g}{M} - (-1)^g \\ &= \frac{(M-1)^{g+1}}{M} \left(1 - \frac{(-1)^g}{(M-1)^g}\right) \\ &= (-1)^{g+1}(M-1) \sum_{j=1}^g \binom{g}{j} (-M)^{j-1}. \end{aligned}$$

*Proof.* ( $g = 1$ ) The element “in the gap” can be chosen in any way except  $y$ , i.e. in  $M - 1$  ways.

( $g = 2$ ) If the first element is chosen in  $M - 1$  ways then the second can be chosen in any way except this way and  $y$ , i.e. in  $M - 2$  ways.

( $g \geq 3$ ) Divide the set of vectors up into disjoint subfamilies according to the first internal occurrence of  $y$ . If there are no occurrences of  $y$  then the number of vectors is  $Q_g(M)$ . Hence

$$\begin{aligned} P_g(M) &= |\{\text{the first internal gap has length 1}\}| \\ &+ |\{\text{the first internal gap has length 2}\}| \\ &+ \cdots \\ &+ |\{\text{the first internal gap has length } g-2\}| \\ &+ |\{\text{there is no internal gap}\}| \\ &= Q_1(M)P_{g-2}(M) + Q_2(M)P_{g-3}(M) + \cdots + Q_{g-2}(M)P_1(M) \\ &+ Q_g(M). \end{aligned}$$

To derive the first form of the closed formula for  $P_g(M)$ , first check explicitly the expressions for  $g = 1$  and for  $g = 2$ . Then use induction and the recurrence (3) to verify the formula for all  $g > 2$ . ■

**DEFINITION 2.3.** Let  $y \in \mathcal{N}$  be given and let  $\mathcal{L} \subset \{1, \dots, N\}$  be a subset. Then  $R(M, N, \mathcal{L})$  is the number of vectors  $x$  in  $M^N$ , satisfying the no-neighbors condition, and  $x_j = y$  for all  $j \in \mathcal{L}$ .

**LEMMA 2.5.** If  $\mathcal{L}$  is empty then  $R(M, N, \mathcal{L}) = M(M-1)^{N-1}$ . If  $\mathcal{L}$  is a singleton,  $\mathcal{L} = \{j\}$ , then  $R(M, N, \mathcal{L}) = (M-1)^{N-1}$ . If  $\mathcal{L} = \{i, j\}$  with  $j - i = g_1 \geq 1$  then

$$R(M, N, \mathcal{L}) = \frac{(M-1)^{N-1}}{M} \left(1 - \left(\frac{-1}{M-1}\right)^{g_1}\right).$$

In general, if  $\mathcal{L}$  has interior gaps  $g_1, \dots, g_{|\mathcal{L}|-1}$ , with  $n_1 \geq 0$  of the  $g_i$  having the value  $g_i = 1$ , then

$$\begin{aligned} R(M, N, \mathcal{L}) &= \frac{(M-1)^{N-1}}{M^{|\mathcal{L}|-1}} \prod_{g_i} \left(1 - \left(\frac{-1}{M-1}\right)^{g_i}\right) \\ &= (M-1)^{N-|\mathcal{L}|} \left(1 - \frac{1}{M}\right)^{|\mathcal{L}|-n_1-1} \prod_{g_i > 1} \left(1 - \left(\frac{-1}{M-1}\right)^{g_i}\right). \end{aligned}$$

*Proof.* The number of vectors associated with each interior gap  $g_i$  is  $P_{g_i}(M)$ . For each end, count vectors from the position of  $y$  “outwards”. If  $g_l$  is the length of the left end then there are associated  $(M-1)^{g_l}$  vectors. If  $g_r$  is the length of the right end then there are associated  $(M-1)^{g_r}$  vectors. The total number of vectors is then found by multiplying the contributions together (since the choices are independent). ■

LEMMA 2.6. Let  $|\mathcal{L}| \geq 2$  and for  $1 \leq i \leq |\mathcal{L}|-1$  let positive integers  $g_i$  be given. Define the polynomial function

$$f(x) = \prod_{g_i} (1 - (-x)^{g_i}) - 1.$$

Then, for  $0 < x < \frac{1}{2}$ ,  $|f(x)| \leq x|\mathcal{L}|(\frac{3}{2})^{|\mathcal{L}|}$ .

*Proof.* Since  $0 < x < 1$ ,

$$\begin{aligned} f(x) &\leq \prod_{g_i} (1+x) - 1 \leq (1+x)^{|\mathcal{L}|} - 1, \text{ and} \\ f(x) &\geq \prod_{g_i} (1-x) - 1 \leq (1-x)^{|\mathcal{L}|} - 1. \end{aligned}$$

Hence

$$|f(x)| \leq \max\{(1+x)^{|\mathcal{L}|} - 1, 1 - ((1-x)^{|\mathcal{L}|})\} = (1+x)^{|\mathcal{L}|} - 1. \quad (1)$$

Let  $g(x) = (1+x)^{|\mathcal{L}|} - 1$ , so  $g(0) = 0$ , and by the Mean Value Theorem, there exists an  $\xi \in (0, x)$  with  $g(x) = g'(\xi)x$ . But  $g'(\xi) = |\mathcal{L}|(1+\xi)^{|\mathcal{L}|-1}$  so, since  $0 < \xi < \frac{1}{2}$ ,

$$\begin{aligned} |g'(\xi)| &\leq x|\mathcal{L}|(\frac{3}{2})^{|\mathcal{L}|} \text{ and therefore by (1)} \\ |f(x)| &\leq x|\mathcal{L}|(\frac{3}{2})^{|\mathcal{L}|}. \end{aligned}$$

In the proof of Theorem 2.1 we use the notation  $\mathcal{L} \subset' \{1, \dots, N\}$ . This means  $\mathcal{L}$  is a subset of  $\{1, \dots, N\}$  which satisfies the no-neighbors condition in the sense that if  $i \in \mathcal{L}$  then neither  $i - 1$  nor  $i + 1$  is in  $\mathcal{L}$ .

**THEOREM 2.1.** *Let  $A'_v(M, N)$  be the average taken over sequences satisfying the no-equal-neighbors condition. If  $M, N \rightarrow \infty$  in such a way that  $N/M \rightarrow \lambda$  with  $0 < \lambda \leq 1$ , then*

$$\lim_{M, N \rightarrow \infty} A'_v(M, N) = \frac{1}{e^\lambda}.$$

*Proof.*

$$\begin{aligned} A'_v(M, N) &= \sum_{\mathcal{L} \subset' \{1, \dots, N\}} (-1)^{|\mathcal{L}|} \frac{(M-1)^{N-1}}{M^{|\mathcal{L}|-1}} \prod_{g_i} \left(1 - \left(\frac{-1}{M-1}\right)^{g_i}\right)^N \\ &\quad \times \frac{1}{M(M-1)^{N-1}N} \\ &= \sum_{\mathcal{L} \subset' \{1, \dots, N\}} \left(\frac{-1}{M}\right)^{|\mathcal{L}|} \prod_{g_i} \left(1 - \left(\frac{-1}{M-1}\right)^{g_i}\right) \\ &= \Sigma_1 + \Sigma_2 \end{aligned}$$

where

$$\begin{aligned} \Sigma_1 &= \sum_{\mathcal{L} \subset' \{1, \dots, N\}} \left(\frac{-1}{M}\right)^{|\mathcal{L}|} \text{ and} \\ \Sigma_2 &= \sum_{\mathcal{L} \subset' \{1, \dots, N\}} \left(\frac{-1}{M}\right)^{|\mathcal{L}|} \left[ \prod_{g_i} \left(1 - \left(\frac{-1}{M-1}\right)^{g_i}\right) - 1 \right]. \end{aligned}$$

Now, by Lemma 2.6, as  $M \rightarrow \infty$ ,

$$\begin{aligned} |\Sigma_2| &\leq \sum_{\mathcal{L} \subset' \{1, \dots, N\}} \left(\frac{1}{M}\right)^{|\mathcal{L}|} \frac{2^{|\mathcal{L}|}}{M} \left(\frac{3}{2}\right)^{|\mathcal{L}|} \\ &= \frac{2}{M} \sum_{j=0}^N \binom{N}{j} \frac{j}{M^j} \left(\frac{3}{2}\right)^j \\ &= \frac{6N}{M^2} \frac{\left(1 + \frac{3}{2M}\right)^N}{2 + \frac{3}{M}}, \end{aligned}$$

which tends to zero as  $M, N \rightarrow \infty$  with  $N/M \rightarrow \lambda$ . Therefore, by Lemma 2.3, we have the main result

$$\lim_{M, N \rightarrow \infty} A'_v(M, N) = \frac{1}{e^\lambda}.$$

■

**COROLLARY 2.1.** *Let  $p$  be an odd prime. Consider the set of sequences of  $p - 1$  components with  $p - 1$  potential values which satisfy the no-equal-neighbors condition. Then the average number of missing values tends to  $1/e$  as  $p \rightarrow \infty$ .*

## REFERENCES

1. Banks, W.D., Luca, F., Shparlinski, I.E. and Stichtenoth, H., *On the value set of  $n!$  modulo a prime*, Turkish J. Math. **29** (2005), p169-174.
2. Cobeli, C., Vâjâitu, M., and Zaharescu, A. *The sequence  $n!(\text{mod } p)$* , J. Ramanujan Math. Soc. **15** (2000), p135-154.
3. Garaev, M. Z., Luca, F., and Shparlinski, I.E. *Exponential sums and congruences with factorials*, J. Reine Angew. Math. **584** (2005), p29-44.
4. Garaev, M. Z., Luca, F., and Shparlinski, I.E. *Character sums and congruences with  $n!$* , Trans. Amer. Math. Soc. **356** (2004), p5089-5102.
5. Guy, R.K., *Unsolved problems in number theory*, Springer, 1994.