THE INNKEEPER'S PROBLEM

MICHAEL RAUGH

Abstract. A tale was once told of a Spanish innkeeper who guessed the probability of randomly placing individual keys for each room on separate hooks in such a way that each key would wind up on the wrong hook. He counted the ways this could occur for an inn with a small number of rooms, and from this small sample he inferred a correct three-figure estimate (≈ 0.368) for his inn of *one hundred rooms*. Here it is explained why his guess was right. In the Afterword I reference two textbook formulations of the same problem, named differently than the innkeeper's problem, and I outline their solution for comparison.

1. The Manuscript

Believe it if you like. Some years ago I was privileged to review an aging manuscript telling a tale about a Spanish innkeeper. The author's signature read "Don Vicente R–y–C de Seville'. Here is the innkeeper's table as recorded by Don Vicente, augmented with my caption.

Rooms	All Ways	Ways for Total Mismatch	Probability
1	1	0	$\uparrow$ 0.0
2	2	1	$\downarrow$ 0.5
3	6	2	$\uparrow$ 0.3333
4	24	9	$\downarrow$ 0.3750
5	120	44	$\uparrow$ 0.3667
6	720	265	$\downarrow$ **0.3681**
7	5040	1854	$\uparrow$ **0.3679**

Table 1. The number of rooms in an inn and, for each number, the various ways room keys can be hung on hooks, and probability of total mismatch. Odd-room probabilities increase, even-room probabilities decrease, squeezing toward $\approx \mathbf{0.368}$. Inkeeper's guess: the same probability must hold for 100 rooms!!!

Date: Feb 8, 2010; Afterword rev 4/11/10.

I will continue to speak of an innkeeper, but of his true identity I have many doubts, and I encourage you to do likewise. But do not doubt the worthiness of the mathematics, which I present in the following sections. In this section I summarize the setting of the problem as given in the manuscript.

Here begins the tale told in the manuscript. With apparently little else to do, the innkeeper with 100 rooms began reckoning. He asked: How likely would it be to get every key on the wrong hook if he placed them randomly?

Table (1) purports to be results from the innkeepers own calculations expressing in the first column the numbers of rooms of an inn, then the total number of "ways" (we would say permutations) for hanging the keys, then the ways (permutations) for hanging keys that leave all keys on the wrong hooks, then in the last column how likely a total mismatch like that would be, or as the results imply, the probability of a total mismatch assuming that all random arrangements are equally likely. I have reported the table to give a flavor of the language of the manuscript. The second column for example refers to factorials.

The manuscript was among linen and other household goods contained in a walnut trunk ornamented with faded painted figures that came over the Santa Fe Trail at an unknown date. This was shown to me by its owner, whose family had long been residents of the old village of Placitas north of the Sandia Mountains in New Mexico. I translated the text from the original Spanish, filled in gaps and made a few corrections, converting formulations into contemporary mathematical notation and methods.

The manuscript was clearly the work of a person of some knowledge of mathematics, perhaps trained by Jesuits as suggested by some Latin phrasing. Apparently, as seen in the table, the "innkeeper", if indeed there was one, knew enough algebra to calculate factorials for the number of ways (permutations) that n keys could be used freely to fill all n slots. And he had to list somewhere the smaller but considerable number of permutations for which no key would be placed in its proper slot. For example, as simple cases, he would have listed for two rooms the one mismatched permutation $(2,1)$, for three rooms the two mismatched permutations $\{(2,3,1),(3,1,2)\}$, for four rooms the nine mismatched permutations $\{(2,1,4,3),(2,3,4,1),(2,4,1,3),\dots\}$, etc. These numbers progress in no obvious pattern, but progress rapidly they do, and that is the problem.

This reckoning grows very tedious as the number of rooms increases, hence the small number of entries in the table. I must suppose the entries in the table for Rooms 6 and 7 (if not the entire table) were completed by Don Vicente based on the mathematics developed in his manuscript. As you can imagine from the magnitude of the factorials in the left-hand column, an actual enumeration would have been entirely too tedious—I say prohibitive—for anyone other than a prodigiously patient savant. If the innkeeper were a savant, I believe Don Vicente would have reported on the unusual capabilities of his host, but there is no such mention.

Another point in evidence is that the table entries are completely accurate—not a single error in any entry. It is scarcely credible that an idle innkeeper could have tabulated all of the totally mismatched permutations, not to mention all of the permutations themselves, without overcounting or undercounting at least one of the entries. I suspect that Don Vicente himself corrected the table, or may even himself created the table, tale and all. But I leave forensics to others.

I now state the surprising fact expressed by the innkeeper that the table implies that for all inns with more than 7 rooms, the probability of mismatching all the keys is approximately 0.368—more rooms would change the result by little. This is an astonishing inference because, though based on few examples, it is correct, and it allows the innkeeper to estimate the probability for 100 rooms. Don Vicente states the innkeeper based his inference on the fact that the right-most entries of the table alternately increase and decrease, with the odd-numbered entries increasing and the even-numbered entries decreasing.[1]

Tragically, the manuscript was lost in a fire that reduced all to rubble—Don Antonio's walnut trunk along with his eighteenth century adobe home—and he has since moved north to Truchas, else would I have provided a scanned image as an appendix. I have retained just enough of my translation to write this tale. I proceed now with the mathematics.

2. The Mathematics of the Manuscript

I begin by following Don Vicente's logic to derive his formulation of the problem dressed in contemporary symbolism. Consider a hotel with $k > 0$ rooms at the inn. There are k numbered keys and k numbered hooks for them behind the front desk.

Let p_k, for an inn of k rooms be the number of permutations for which *none* of the k keys is on the right hook. Define $p_0 = 1$.[2] Then for an inn of n rooms we can count the number of such permutations in

[1] These tendencies I have noted with arrows in the table for ease of reading.

[2] This is like defining $0! = 1$ to simplify some expressions.

terms of the number of permutations for inns with fewer rooms, giving us an equation called a *recusion* formula:

$$p_n = n! - \sum_{k=0}^{n-1} \binom{n}{k} p_{n-k} \tag{1}$$

Eq. (1) is called a recursion formula because it determines p_n recursively in terms of earlier terms p_k. The idea in Eq. (1) is that if exactly k keys are distributed correctly, then exactly p_{n-k} keys are distributed incorrectly. There are $\binom{n}{k}$ ways to choose k rooms out of n rooms, so there must be $\binom{n}{k} p_{n-k}$ ways to get exactly k keys on the right hook.

In order to count the number of permutations of keys for which all are hung *incorrectly*, we must subtract from $n!$ all the permutations for which at least one key is placed *correctly*. Eq. (1) does this.

We can simplify Eq. (1) by moving the summation over to the left and absorbing p_n into it:

$$\sum_{k=0}^{n} \binom{n}{k} p_{n-k} = n!$$

Or even better, here's an equivalent formulation obtained by dividing $n!$ out of all the binomial coefficients:

$$\sum_{k=0}^{n} \frac{p_{n-k}}{(n-k)!} \cdot \frac{1}{k!} = \sum_{k=0}^{n} \frac{p_k}{k!} \cdot \frac{1}{(n-k)!} = 1 \tag{2}$$

This is the form of the recursion formula we will use. We have an equation that determines p_n recursively, but we would like to have a simple formula for p_n just in terms of n. In the following sub-section we find such a formula using a technique pioneered by Euler.

Eq. (1) is in essence the formulation achieved by Don Vicente. We recognize it as a recursion formula for which solution methods exist. But Don Vicente instead worked out progressive solutions for $p_1, p_2, \cdots, p_{10}$ laboriously, and from these surmised a general solution, a creditable achievement.

2.1. **An Eulerian Solution of the Recursion Formula.** The expression under summation in Eq. (2) is suggestive of the coefficients that arise in multiplying two power series. To see why, we need to do the kind of thing Euler thought of. Define a function $P(x)$ by incorporating the numbers p_k into a power series,[3]

$$P(x) = \sum_{n=0}^{\infty} \frac{p_n}{n!} x^n$$

We'll have to figure out a better way to express this function. But first let's see how the following equation gives us our recursion formula.

$$\left(\sum_{n=0}^{\infty} \frac{p_n}{n!} x^n\right) \left(\sum_{n=0}^{\infty} \frac{x^n}{n!}\right) = \sum_{n=0}^{\infty} x^n \tag{3}$$

We take for granted *Cauchy's product rule* for infinite series, which says that we can multiply infinite series as though they were ordinary polynomials, beginning with the constant terms and working through increasing powers of x. For example, the term for x^3 in the product on the left-hand side of Eq. (3) works out to be,

$$\left(\frac{p_0}{0!}\frac{1}{3!} + \frac{p_1}{1!}\frac{1}{2!} + \frac{p_2}{2!}\frac{1}{1!} + \frac{p_3}{3!}\frac{1}{0!}\right) x^3$$

Equating this coefficient with the coefficient of x^3 on the right-hand side of Eq. (3) yields the recursion formula for the case for $n = 3$ expressed by Eq. (2). The same idea works in general: equating the coefficients of x^n on the left- and right-hand sides of Eq. (3) for any value of n gives us the recursion formula of Eq. (2) for that same value of n.

The second parenthesis in Eq. (3) is the power series for e^x. So we can re-write that equation:[4]

$$\sum_{n=0}^{\infty} \frac{p_n}{n!} x^n = e^{-x} \sum_{n=0}^{\infty} x^n \tag{4}$$

$$= \left(1 - \frac{x}{1!} + \frac{x^2}{2!} - \frac{x^3}{3!} + \cdots\right) \left(1 + x + x^2 + x^3 + \cdots\right)$$

[3]The series converges on at least the interval $(-1, 1)$ because $|p_n| \leq n!$.

[4]The operations are justified because the series all converge on the common interval $(-1, 1)$, though of course they are not all restricted to that interval.

The function expressed on the right-hand sides of Eq. (4) is called a *generating function* for the coefficients on the left-hand side. It reduces solving the recursion formula Eq. (2) to magical simplicity, as follows.

The product of the series on the right-hand side of Eq. (4) yields,

$$\sum_{n=0}^{\infty} \frac{p_n}{n!} x^n = \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} (-1)^k \frac{1}{k!} \right) x^n$$

Therefore, by equating coefficients of equal powers of x on both sides of the equation, we find that,

$$\frac{p_n}{n!} = \sum_{k=0}^{n} (-1)^k \frac{1}{k!} \tag{5}$$

Since $\dfrac{p_n}{n!}$ is the probability of getting all n keys on the wrong hook—and since the alternating series converges *very* rapidly to $1/e$—we see that the probability converges to $1/e \approx \mathbf{0.368}$ as $n \to \infty$, the same number inferred by the innkeeper.

3. Afterword

I have left it to you to decide what's true and false in the tale of the innkeeper. But that there is a problem of mathematical interest exposed therein, there can be little doubt.

According to Feller [1, 100], the problem has many variants going back to a card-matching problem of Montmart of 1708. I have also seen the problem as one of mismatching letters and envelopes. I first learned of the problem as one of mismatching the first n integers with numbered slots in Courant and Robbin's *What Is Mathematics? An Elementary Approach to Ideas and Methods* [2, 114–116]. I refer to the recent edition augmented by Ian Stewart.

In essence, the idea is to count without redundancy the number of *unique* elements in the union of a finite number of sets. The complication is that any particular element of the union may belong to more than one of the sets, so overcounting has to be nullified. By formulating the problem in terms of probability, the proof in *What Is Mathematics?* is the more straightforward. Let $P(E)$ be the probability of an event E. Courant and Robbins begin with the equation,

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

which follows from the fact that $P(A)+P(B)$ counts $P(A\cap B)$ twice. Substituting the expression $B\cup C$ for B in the equation above and expanding using the equation again yields,

$$P(A\cup B\cup C) \tag{6}$$
$$= P(A)+P(B)+P(C)$$
$$-P(A\cap B)-P(A\cap C)-P(B\cap C)$$
$$+P(A\cap B\cap C)$$

The emerging pattern for generalization to an arbitrary number of sets $A,B,C,D,\ldots$ is easily discovered, and Courant and Robbins leave it as an exercise.

Courant and Johns' approach (which see) is equivalent to Feller's, but Feller's is expressed in general terms, and his method of counting is explicit. I follow Feller here.

Let $A_1,A_2,\ldots,A_n$ be a collection of subsets of some larger set. Suppose a is an element of the union $\bigcup_{j=1}^{n} A_j$. If a belongs to *exactly* k of the sets, then it appears k times among the sets $A_1,A_2,\ldots,A_n$, and it appears $\binom{k}{2}$ times among the intersections of *pairs* of those sets, and it also appears $\binom{k}{3}$ times among the intersections of *triplets* of those sets, and so forth. The trick in the method is to play these numbers off against one another using the binomial theorem.[5]

To state this precisely, let $\mathfrak{A}^{(j)}, j=1,\ldots,n$ represent the set of all j-way intersections among the sets $A_1,A_2,\ldots,A_n$. Then a appears in exactly $\binom{k}{j}$ of the sets of $\mathfrak{A}^{(j)}$, for $j=1,\ldots,k$, but a does *not* appear in any of the sets $\mathfrak{A}^{(j)}$, for $j>k$.

Now we play the numbers against one another. Consider the alternating sum,[6]

$$S(A_1,A_2,\ldots,A_n)=\sum_{j=1}^{n}(-1)^{j-1}\bar{\mathfrak{A}}^{(j)} \tag{7}$$

[5]Use the fact that $(1-1)^k=\sum_{i=0}^{k}(-1)^j\binom{k}{j}=0$.

[6]This is a combinatoric equivalent to the generalization of Eq. (6).

where $\bar{\mathfrak{A}}^{(j)}$ denotes the combined sum of the numbers of elements in all the sets of $\mathfrak{A}^{(j)}$, a sum that will count the element a exactly $\binom{k}{j}$ times for $j = 1, \ldots, k$, and not at all for $j > k$. Therefore, what a contributes to $S(A_1, A_2, \ldots, A_n)$ is,

$$\sum_{j=1}^{k} (-1)^{j-1} \binom{k}{j} = 1 - (1-1)^k = 1$$

Since a represents a typical element of the union $\bigcup_{j=1}^{n} A_j$, and it gets counted only once in $S(A_1, A_2, \ldots, A_n)$, it follows that $S(A_1, A_2, \ldots, A_n)$ counts each element of the union only once. This proves that Eq. (7) cancels all overcounting and counts *exactly* the number of elements in $\bigcup_{j=1}^{n} A_j$.

This method of counting the number of elements in a union of sets, some of which may be overlapping, is not so obvious. But once we have the counting formula in hand, it can be applied to solve the innkeeper's problem.

Solving inkeeper's problem by counting. Let $A_j, j = 1, \ldots, n$ represent the set of *all* permutations of the set of integers $1, \ldots, n$ such that the integer j appears in position j. This is analogous to the set of all permutations of n keys such that at least the j-th key is correctly hung on the j-th hook.

Note that for a permutation to be in A_j, we require *only* that the integer j appear in the j-th position—we do *not* care how the other integers are permuted. If $n > 1$, some permutations in A_j will have any of the integers other than j placed correctly, too. For example, A_j contains the permutation with all of the integers $1, \ldots, n$ placed correctly.

Let $\mathfrak{A}^{(m)}$ represent all the m-way intersections of sets from among the sets $A_1, A_2, \ldots, A_n$. Such an intersection can be specified by $A_{j_1} \cap \cdots \cap A_{j_m}$, where $\{j_1, \ldots, j_m\}$ is a selection of m distinct integers from among the first n positive integers. This intersection certainly is not empty since by definition each of the sets A_{j_k} contains the permutations in which all numbers $\{j_1, \ldots, j_m\}$ are place correctly, and we can count how many such permutations there are: altogether there are $(n-m)!$, since the remaining integers among $1, \ldots, m$ can be permuted arbitrarily.

Moreover, there are $\binom{n}{m}$ ways of selecting m-way intersections of sets from among the sets $A_1, A_2, \ldots, A_n$. So the combined sum of all of the elements in these various intersections, is

$$\bar{\mathfrak{A}}^{(m)} = (n-m)!\binom{n}{m} = \frac{n!}{m!}$$

Sustituting this result in Eq. (7) tells us that the number S of ways that at least one integer can be placed correctly is,

$$S = n!\sum_{m=1}^{n}(-1)^{m-1}\frac{1}{m!}$$

Because there are $n!$ possible permutation of the integers $1, \ldots, n$, and we may suppose that all the permutations are equally likely, the *probability* of at least one integer being placed correctly is,

$$\frac{S}{n!} = \sum_{m=1}^{n}(-1)^{m-1}\frac{1}{m!}$$

Taking the complement shows that the probability of *no integer* being placed correctly is,

$$\frac{p_n}{n!} = 1 - \frac{S}{n!} = \sum_{m=0}^{n}(-1)^{m}\frac{1}{m!}$$

Interpreting this in terms of the innkeeper's problem—keys instead of integers—we see that this is the same result as in Eq. (5).

You may now compare the textbook solutions of the total-mismatch problem—named here the innkeeper's problem—with the one I have given based on Don Vicente's recursion formula. There is something to be said for each approach. In my opinion, the method used in Section (2.1) arises step-by-step more naturally than the textbook solutions, if you are familiar with analysis. In addition, the method of generating functions reminiscent of Euler has wide applicability in analysis, so that method is perhaps more illustrative of a general technique. And, remarkably, it is an example of the use of analysis to produce a discrete result about permutations. What's your opinion?

References

1. William Feller, *An introduction to probability theory and its applications*, 3 ed., vol. 1, John Wiley & Sons, Inc., New York, 1968.
2. Richard Courant, Herbert Robbins and Ian Stewart, *What is mathematics? an elementary approach to ideas and methods*, Oxford University Press, USA, 1996.
3. C. Edward Sandifer, Editor, *The early mathematics of leonhard euler*, The Mathematical Association of America, Washington, DC, 2007.

E-mail address: `michael dot raugh (at) gmail dot com`
URL: `www.mikeraugh.org`