HOW LONG UNTIL A RANDOM SEQUENCE DECREASES?

JACOB A. SIEHLER

Abstract. We consider the position of the first decrease in a sequence of numbers generated by a random variable \( X \), and prove simple formulas for the expected value of this position when \( X \) is a uniformly distributed discrete variable or an arbitrary continuous random variable. Specifically, in the continuous case, we show that expected value is \( e \), independent of the distribution of \( X \).

1. Waiting for the fall

Imagine observing a stream of random real numbers: if you saw the sequence

\[
0.0478, 0.1429, 0.1667, 0.2204, 0.8124, 0.8226, 0.3101 \ldots,
\]

with the first decrease occurring in the seventh position, you might feel that this was an unusually long time to wait for that first decrease – even if you’re not exactly sure how long such a run “usually” lasts in a random sequence. The average position of the first decrease in a stream of random numbers depends on precisely what you mean by random; that is, it depends on the distribution of random numbers that you’re sampling. Surprisingly, though, for \( \textit{continuous} \) distributions, the question has a very specific (and delightful) answer that is independent of how the random numbers are distributed.

We will uncover the answer in due time – in Proposition 7, to be precise (it’s not a bad idea to exercise your intuition by making a guess in advance about how long, on average, a monotone run like the one above will last in a sequence of random numbers)\(^1\). Along the way, we’ll get a clinic in the interaction between calculus and discrete math, with derivatives, famous limits, power series, and the binomial theorem playing key roles. We can start by discussing a special case of the problem, when the random numbers are generated simply by rolling dice.

2. The die-rolling game

One of the fringe benefits of teaching a course on probability and statistics is that it affords an excellent excuse to keep an assortment of toys on my desk, especially all sorts of dice. This article had its beginning when I was rolling an ordinary 6-sided die and got what I felt was an unusually long run before the first decrease in the numbers occurred. It might have been something like this:

\[
1, 3, 3, 3, 4, 6, 5
\]

and I decided to make a game of it: I would award myself 7 points for that run, since I got to roll seven times (including the final, decreasing roll that ended the game). Naturally, I was wondering what a typical score in my new game might be. But I also had my eye on

\(^1\)If you’ve been around math much, you can probably make a shrewd guess based solely on the fact that I described the answer as “delightful.”
the 4- and 20-sided dice lying nearby, and wondered if one of those might give me a better chance to get a large score. (Yes, this is another good point to exercise your intuition: can I expect to get a better score by choosing one of the other dice?)

Mathematically, an $n$-sided die is modelled by a discrete random variable that is uniformly distributed over the set $\{1, 2, \ldots, n\}$ – that is, all of the outcomes in the set occur with the same probability, $1/n$. But we’re not interested in single rolls of the die; rather, we want to study non-decreasing runs from repeated rolls of the die, and that motivates the following:

**Definition 1.** Given any random variable $X$ we define an associated random variable $R(X)$, the run-length variable for $X$, as follows: we sample $X$ until the first decrease occurs, then let $R(X)$ be the total number of samples we took, including the final decrease that ends the experiment. Call $X$ the underlying variable of $R(X)$.

The $R(X)$ notation emphasizes the fact that the experiment depends on the underlying variable $X$, but we’ll suppress the subscript and simply refer to $R$ when the underlying variable is clear from context. And we’ll call the run-length variable $R_n$ when $X$ is an $n$-sided die roll.

Our goal now is to study the expected value of $R$, particularly when the underlying variable is an $n$-sided die roll. Informally, the expected value of a random variable is the long-term average of its outcomes; by definition, the expected value of $R$ is

$$E[R] = \sum_r r f(r)$$

where the summation is over all possible outcomes $r$ that might occur, and $f(r)$ denotes the probability of getting outcome $r$.

No matter what the underlying variable is, we always get an outcome of at least 2 for $R$. On the other hand, there’s no upper bound on the potential length of a nondecreasing run, so the possible outcomes of $R$ are $\{2, 3, 4, \ldots\}$, and we can rewrite the expected value calculation with more explicit limits of summation:

$$E[R] = \sum_{r=2}^{\infty} r f(r)$$

The first mathematical hurdle in understanding the die-rolling game, therefore, is determining a formula for the function $f(r)$ (the “probability mass function”) which tells how likely an outcome of $r$ is in our game.

**Proposition 1.** For $n \geq 2$, the probability mass function for $R_n$ (run lengths for an $n$-sided die roll) is given by

$$f_n(r) = \frac{(n+r-2)(r-1)}{n^r}$$

for all $r \geq 2$.

**Proof.** Let $a_r$ denote the number of nondecreasing sequences of length $r$ that can be formed from the set $\{1, 2, \ldots, n\}$. This is a problem of selection with repetition allowed, and a standard result from combinatorics [5] is that $a_r$ is given by a binomial coefficient:

$$a_r = \binom{n+r-1}{n-1}$$

The number of runs of length $r$ is

$$a_r = \binom{n+r-2}{r-1}$$

The probability of getting a run of length $r$ is

$$f_n(r) = \frac{a_r}{n^r} = \frac{(n+r-2)(r-1)}{n^r}$$
Imagine we have already established a nondecreasing run of length \( r \), and we want to know the probability that we’ll succeed at extending this to a nondecreasing run of length \( r + 1 \) with our next roll. Call this probability \( S_n(r) \) (success after the \( r \)th roll). We might have any of the \( a_r \) nondecreasing sequences in hand, and there are \( n \) ways in which we could extend our sequence – but there are only \( a_{r+1} \) good sequences of length \( r + 1 \). So we have

\[
S_n(r) = \frac{a_{r+1}}{na_r} = \frac{n+r}{n\binom{n+r-1}{n-1}}
\]

which, after merciless cancellation of factorials, simplifies to

(2) \[
S_n(r) = \frac{n + r}{n(r + 1)}
\]

Now, to get an outcome of precisely \( r \) rolls, we need to succeed at extending \( r - 2 \) times after our first roll, then finally fail to extend after the \( (r - 1) \)th roll. Since successive rolls of the die are independent of one another, that means

\[
f_n(r) = S_n(1) \times S_n(2) \times \cdots \times S_n(r - 2) \times (1 - S_n(r - 1))
\]

so we can use our formula (2) and simplify:

\[
f(r) = \frac{(n + 1)(n + 2)\ldots(n + r - 2)}{n^{r-2}(r - 1)!} \left(1 - \frac{n + r - 1}{nr}\right)
\]

\[
= \frac{(n + r - 1)!}{r!(n - 2)!n^r}
\]

\[
= \frac{(n + r - 2)(r - 1)}{n^r}
\]

as claimed. \( \square \)

With a formula for the mass function established, we can make an assault on the problem of determining the expected value of \( R_n \).

**Proposition 2.** The expected value of \( R_n \) given exactly by

\[
E[R_n] = \left(\frac{n}{n-1}\right)^n
\]

**Proof.** We start with nothing more sophisticated than substituting formula (1) into the expected value summation, and hope for the best:

\[
E[R_n] = \sum_{r=2}^{\infty} rf_n(r) = \sum_{r=2}^{\infty} r \cdot \frac{(n+r-2)(r-1)}{n^r}
\]

That can be put in a slightly more helpful form by rewriting the binomial coefficient in terms of factorials.

\[
E[R_n] = \sum_{r=2}^{\infty} \frac{(r-1)(r)(r+1)\ldots(r+n-2)}{(n-2)!} \left(\frac{1}{n}\right)^r
\]

On the whole, it doesn’t look terrifically appealing, but the \((1/n)^r\) term suggests that we might be looking at a power series in, say, \( x \), which has been evaluated at \( x = (1/n) \). If we forget about this specialized value of \( x \) for a moment, we can use calculus to help with the
summation – and doesn’t that pile of coefficients up front look like the sort of thing that might occur due to repeated differentiation? Replacing \((1/n)\) with a generic \(x\), we have

\[
\sum_{r=2}^{\infty} \frac{(r-1) \cdots (r+n-2)x^r}{(n-2)!} = \frac{x^2}{(n-2)!} \sum_{r=2}^{\infty} (r-1) \cdots (r+n-2)x^{r-2}
\]

\[
= \frac{x^2}{(n-2)!} \frac{d^n}{dx^n} \left[ \sum_{r=2}^{\infty} x^{n+r-2} \right]
\]

(\(*\))

Now we’re looking at a geometric series with ratio \(x\), whose first term is \(x^n\), and anybody\(^2\) can sum that:

\[
(\ast) = \frac{x^2}{(n-2)!} \frac{d^n}{dx^n} \left[ \frac{x^n}{1-x} \right],
\]

with the series converging if and only if \(|x| < 1\).

Calculus gives and takes away; we have disposed of the summation, but we have to find a nice way to apply differentiation repeatedly to the expression in brackets. A standard helpful move for differentiating rational functions is to use long division to reduce to a polynomial and a proper fraction. In this case, after applying long division to \(x^n/(1 - x)\) we get an expression that’s very easy to differentiate.

\[
(\ast) = \frac{x^2}{(n-2)!} \frac{d^n}{dx^n} \left[ -(1 + x + \cdots + x^{n-1}) + \frac{1}{1-x} \right]
\]

\[
= \frac{x^2}{(n-2)!} \cdot \frac{n!}{(1-x)^{n+1}}
\]

\[
= \frac{x^2 n(n-1)}{(1-x)^{n+1}}
\]

Remember, we were particularly interested in the value of this sum at \(x = (1/n)\). Since we have assumed \(n \geq 2\), the series does converge, and evaluating the preceding expression at \(x = (1/n)\) yields

\[
E[R_n] = \frac{(1/n)(n-1)}{(1-1/n)^{n+1}} = \left( \frac{n}{n-1} \right)^n,
\]

completing the proof of the proposition.

Let’s look back at a few questions we can now answer about the die-rolling game:

1. By Proposition 1, the probability that I would get a score of 7, using a 6-sided die as in the example, is just

\[
f(7) = \left( \frac{11}{6} \right) \cdot 6
\]

which is about 0.007 – small enough that you might suspect my example is fictitious. (With a little more work you can check that the probability that I would get a score of \( \text{at least} 7 \) is just a tiny bit less than 1%, which is probably a more relevant fact.)

2. By Proposition 2, the average score for a game with a 6-sided die would be \((6/5)^6\), or just barely under 3.

\(^2\)Even the author, which really proves the point.
(3) Since the formula \( \left( \frac{n}{n-1} \right)^n \) is strictly decreasing in \( n \), I’d have a higher score, on average, if I switched to a 4-sided die, and a lower score if I used the 20-sided die. In fact, to maximize your score, your best bet for this game would be to toss a coin with sides labelled 1 and 2. You’d have an expected score of 4 in that case.

If you’re still thinking about the question posed in the introduction, you might stop to consider this: What can you say about the expected value of \( R \) if \( X \) has a very large number of equally likely outcomes?

3. A variation: strictly increasing die rolls

If we change the rules of the die-rolling game slightly to insist on strictly increasing numbers, we get slightly different results for the mass function and expected score. In the strictly increasing game, for example, rolling

\[ 1, 3, 5, 5 \]

would cause the game to end with a score of 4. We can briefly establish results analogous to those of the previous section for the strictly increasing game. The expected value calculation has a rather different flavor to it, which is an interesting contrast.

**Definition 2.** Given any random variable \( X \), we define another random variable \( R^s(X) \) as follows: we sample \( X \) until we get a result which is not strictly greater than the previous result, then let \( R^s(X) \) be the total number of samples we took.

**Proposition 3.** If \( X \) represents an \( n \)-sided die roll then the probability mass function for \( R^s(X) \) is given by

\[
f^s(r) = \frac{\binom{n+1}{r} (r-1)}{n^r}
\]

The proof is left as an exercise; it is very similar to the proof of Proposition 1.

**Proposition 4.** If \( X \) represents an \( n \)-sided die roll then the expected value of \( R^s(X) \) is given by

\[ E[R^s(X)] = \left( \frac{n+1}{n} \right)^n \]

**Proof.** Notice that in this variant we have an upper bound on the possible scores: if we use an \( n \)-sided die, then our score must come from the set \( \{2, \ldots, n+1\} \). That means that the expected value calculation involves only a finite sum; using the preceding proposition, we have

\[
E[R^s] = \sum_{r=2}^{n+1} r \cdot \frac{\binom{n+1}{r} (r-1)}{n^r}
\]

which we can rewrite as

\[
E[R^s] = \sum_{r=2}^{n+1} \frac{(n+1)!r(r-1)}{r!(n-r+1)!n^r}
\]
This becomes more tractable if we reindex; let \( j = r - 2 \), so \( r = j + 2 \). Then we can write

\[
E[R^s] = \sum_{j=0}^{n-1} \frac{(n+1)!}{(j+2)!(n-j-1)!} \frac{n^{j+2}}{(j+1)^j}
\]

\[
= \left( \frac{n+1}{n} \right) \sum_{j=0}^{n-1} \frac{(n-1)!}{j!(n-1-j)!} \left( \frac{1}{n} \right)^j
\]

\[
= \left( \frac{n+1}{n} \right) \sum_{j=0}^{n-1} \binom{n-1}{j} \left( \frac{1}{n} \right)^j
\]

and this last summation is exactly the expansion of \((1 + \frac{1}{n})^{n-1}\) given by the binomial theorem. So,

\[
E[R^s] = \left( \frac{n+1}{n} \right) \left( 1 + \frac{1}{n} \right)^{n-1} = \left( \frac{n+1}{n} \right)^n
\]
as claimed. \( \square \)

**Question.** In the strictly increasing game, what sort of die (i.e., how many sides) should you choose to maximize your expected score?

## 4. The Continuous Game

Now, instead of rolling dice to generate our random numbers, suppose we have a continuous (real) random variable \( X \) as our source of randomness.

The only assumption we will make is that \( X \) is described by a probability density function\(^3\) – that is, there is a nonnegative function \( p(x) \) on \( \mathbb{R} \) with the property that the probability of getting an outcome from \( X \) that is between \( a \) and \( b \) is given by

\[
P(a \leq X \leq b) = \int_a^b p(x) \, dx
\]

Graphically, that probability is represented by the area bounded by the graph of \( p(x) \) between \( a \) and \( b \). Figure 1 shows the familiar bell shape of the normal distribution, and the density function for a more fanciful distribution.

![Figure 1](image)

**Figure 1.** A normal and a not-so-normal distribution, with \( P(a \leq X \leq b) \) represented by areas on the graph of the density function.

The variable \( R(X) \) from Definition 1 works just the same as in the discrete case, and we’ll refer to the process of sampling \( X \) to determine \( R(X) \) as “playing the continuous game.”

---

\(^3\)The density function itself does not need to be continuous; its integrals will be, and that’s all we need.
To understand the continuous game, we can make use of our earlier results on die-rolling games if we consider a discretized version of the continuous game: fix an \( n \geq 2 \), and cut the distribution for \( X \) into \( n \) equally likely quantiles, numbered 1 to \( n \). For example, Figure 2 shows a distribution cut into six equally likely quantiles; each of the six pieces (including the two tails) contains the same amount of area.

![Figure 2](image)

**Figure 2.** A sequence of samples on a probability distribution cut into equally likely quantiles.

If we play the continuous game and record a sequence of reals

\[
x_1, x_2, x_3, x_4, x_5, x_6
\]

as shown in Figure 2, then in the discretized version, we would record instead the sequence

1, 1, 4, 4, 6, 5

showing only the numbers of the quantiles our samples landed in.

To set up the next lemma, suppose we have already established a run

\[
x_1 \leq x_2 \leq \cdots \leq x_r
\]

in the continuous game and let \( S(r) \) denote the probability of successfully extending to a run of length \( r + 1 \). As in Proposition 1, \( S_n(r) \) denotes the probability of getting a result that extends our run in the discretized version.

**Lemma 5.** \( S(r) = \lim_{n \to \infty} S_n(r) \).

**Proof.** Any outcome that extends our run in the continuous game also extends our run in the discretized game, so

\[
(4) \quad S(r) \leq S_n(r)
\]

On the other hand, the only way we can succeed at extending in the discretized game without succeeding in the continuous game is to get a result smaller than \( x_r \) but in the same quantile. Since the probability of getting a result anywhere in that quantile is only \( 1/n \), we have

\[
(5) \quad S_n(r) \leq S(r) + (1/n).
\]

Applying the Squeeze Theorem to (4) and (5) completes the proof. \( \square \)
Since the probability mass function \( f_n(r) \) for the discretized game is determined by the \( S_n(r) \)'s, and we have a formula for \( f_n(r) \), we can obtain an equally nice formula for the mass function in the continuous game.

**Proposition 6.** If \( X \) is any continuous random variable, then the probability mass function for the run-length variable \( R(X) \) is given by

\[
f(r) = \frac{r - 1}{r!}
\]

for \( r \geq 2 \).

**Proof.** Fix any \( r \geq 2 \). As in Proposition 1, we have

\[
f(r) = S(1) \times S(2) \times \cdots \times S(r - 2) \times (1 - S(r - 1))
\]

but by the previous lemma, each \( S \) on the right-hand side of this expression is a limit of \( S_n \)'s. That means that \( f(r) \) is a limit of \( f_n(r) \)'s:

\[
f(r) = \lim_{n \to \infty} S_n(1) \times S_n(2) \times \cdots \times S_n(r - 2) \times (1 - S_n(r - 1))
\]

\[
= \lim_{n \to \infty} f_n(r)
\]

If we use our formula (1) for \( f_n(r) \), we can take the limit algebraically and obtain the formula for \( f(r) \).

\[
f(r) = \lim_{n \to \infty} \frac{(n + r - 2)(r - 1)}{n^r}
\]

\[
= \lim_{n \to \infty} \frac{(n - 2 + r)! (r - 1)}{(n - 2)! r! n^r}
\]

\[
= \lim_{n \to \infty} \frac{r - 1}{r!} \left[ \frac{(n - 2 + r) \times \cdots \times (n - 2 + 1)}{n^r} \right]
\]

\[
= \frac{r - 1}{r!}
\]

since the numerator and denominator of the bracketed expression are both monic polynomials of degree \( r \) in \( n \). \( \square \)

**Proposition 7.** Let \( R \) be the run-length variable for any continuous random variable \( X \). Then the expected value of \( R \) is exactly

\[E[R] = e,\]

the base of the natural logarithm.

**Proof.** This is really an easy corollary of the preceding proposition; starting from the definition of expected value we have

\[
E[R] = \sum_{r=2}^{\infty} r f(r) = \sum_{r=2}^{\infty} \frac{r(r - 1)}{r!} = \sum_{r=2}^{\infty} \frac{1}{(r - 2)!}
\]

and this last expression is exactly the beloved series expansion

\[1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \cdots = e.\] \( \square \)
Remark. We could have arrived at this result heuristically as follows: a continuous distribution is, loosely speaking, like a distribution with infinitely many equally likely outcomes. Since the expected run length when there are \( n \) equally likely outcomes is \( E[R_n] = (n/(n-1))^n \), we could have guessed that 

\[
E[R] = \lim_{n \to \infty} \left( \frac{n}{n-1} \right)^n
\]

and this is another famous limit expression for \( e \). (We could have used a limit of formula (3) from the strictly increasing game, too.) Modelling a proof directly on this argument is not impossible, but has some difficulties; moreover, it doesn’t reveal the simple formula for the mass function in the continuous game.

The appearance of \( e \) in this problem is reminiscent of its appearance in the “Hat-Check Problem” (see [2] or [3]), where \( 1/e \) occurs as the approximate probability that a random permutation of an \( n \)-element set has no fixed points; the probabilities converge to \( e \) as \( n \) gets large. In the hat-check problem, \( 1/e \) is an excellent approximation to the true probability even for relatively small \( n \). In our problem, the expected value for \( R_n \) converges much more slowly to \( e \); roughly, you have to use an \( n = 10^k \)-sided die for \( E[R_n] \) to match \( k \) decimal digits of \( e \). The slow convergence of this sequence is discussed in [4].

5. A combinatorial connection

The discrete distribution described by Proposition 6 can be viewed as a limiting case of either of the two families of distributions introduced in §2 and §3. Neither the families nor the limiting distribution appear to be familiar enough to have a widely-known name attached to them. For the statistically inclined and curious, further investigation of these distributions might begin with their variance and higher moments. In the case of the run-length distribution for continuous variables, that would entail considering sums of the form

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

for different exponents \( k \). Starting with \( k = 1 \), this will yield a sequence beginning

\[
e, 3e, 10e, 37e, 151e, 674e, \ldots
\]

The sequence of coefficients (A005493 in [7]) has a combinatorial interpretation in its own right, but may be better recognized as first differences in the sequence of Bell numbers,

\[
\{B_k\} = 1, 2, 5, 15, 52, 203, 877, \ldots,
\]

which count the number of ways to partition a \( k \)-element set. The connection can be seen by splitting (6) as

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

and recognizing these as Dobinski’s summations (see [6] or [1]) for \( B_{k+1} \) and \( B_k \) (albeit with their first terms deleted).

\[\text{Or stumbling across them in the literature.}\]
REFERENCES


DEPARTMENT OF MATHEMATICS, WASHINGTON & LEE UNIVERSITY, LEXINGTON, VA 24450

E-mail address: siehlerj@wlu.edu