The Surprising Predictability
Author(s): Mark F. Schilling
Reviewed work(s):
Source: *Mathematics Magazine*, Vol. 85, No. 2 (April 2012), pp. 141-149
Published by: Mathematical Association of America
Stable URL: http://www.jstor.org/stable/10.4169/math.mag.85.2.141
Accessed: 14/10/2012 13:35

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at
http://www.jstor.org/page/info/about/policies/terms.jsp

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Mathematical Association of America is collaborating with JSTOR to digitize, preserve and extend access to *Mathematics Magazine*.

http://www.jstor.org

to prove some standard results. The advantage of the "M-method" is to avoid having to estimate the remainder term. Both approaches, though, should be presented in a course.

Addendum. After a first revised version of this note had been submitted, I learned from Professor Ingo Lieb that the notion that I call M-differentiability is not new; and is, in fact, contained in several German textbooks (see [**1, 2, 3, 4**]). It was mainly popularized by Professor Hans Grauert and its school. Some historical aspects on this approach appear also in [**5**] and [**6**].

Acknowledgment I warmly thank Professor Ingo Lieb and Professor Michael Range for their remarks and for forwarding to me the references below.

REFERENCES

1. W. Fischer and I. Lieb, *Einführung in die Komplexe Analysis*, Vieweg + Teubner, Wiesbaden, 2010.
2. H. Grauert and W. Fischer, *Differential-und Integralrechnung II*, Springer, Berlin, 1978.
3. H. Grauert, W. Fischer, and I. Lieb, Differential and integral calculus, *Izdat. "Mir,"* Moscow, 1971 (in Russian).
4. H. Grauert and I. Lieb, *Differential-und Integralrechnung I*, Springer, Berlin, 1973.
5. E. Hairer and G. Wanner, *Analysis in historischer Entwicklung*, Springer, Berlin, 2010.
6. M. Range, Where are limits needed in calculus? *Amer. Math. Monthly* **118** (2011) 404–417; http://dx.doi.org/10.4169/amer.math.monthly.118.05.404.

Summary Based on the notion of M-differentiability, we present a short proof of the differentiability of composite functions in the finite dimensional setting.

The Surprising Predictability of Long Runs

MARK F. SCHILLING
California State University, Northridge
Northridge, CA 91330
mark.schilling@csun.edu

On August 18, 1913, a growing crowd gathered around a roulette table in Monte Carlo as the wheel stopped at black numbers on 26 consecutive spins. On May 15, 1941, baseball player Joe DiMaggio began a hitting streak that lasted for 56 games. And the state of Missouri voted for the winning candidate in every presidential election from 1904 to 1952, then did it again from 1960 to 2004. Long runs of identical outcomes often attract considerable attention, and seem to call for explanation. While any of these events is remarkable in isolation, it turns out that their existence, and especially their lengths, are surprisingly predictable.

When a situation can be modeled as a sequence of independent Bernoulli trials, a simple rule of thumb predicts the length L of the longest run of successes, often with remarkable accuracy. Sometimes we can practically guarantee the exact value of L. The key result is the approximation of the distribution of L by an extreme value distribution. We provide applications to several situations, including the California State Lottery and the digits of π.

Math. Mag. **85** (2012) 141–149. doi:10.4169/math.mag.85.2.141. © Mathematical Association of America

The approximate length of the longest success run

Suppose we have a sequence of n independent Bernoulli trials, each with success probability p. (The classic example is a fair coin tossed n times, with success being the outcome "heads" and $p = 1/2$.) By a "success run" we mean a sequence of one or more consecutive successes at the start of the sequence or after any failure. We define the random variable $L = L_n$ to be the length of the longest success run. If N_F is the random variable representing the number of failures, then $E(N_F) = nq$, where $q = 1 - p$, and so we should expect about nq opportunities for success runs. (We assume that $nq \gg 1$, and take that as license to disregard any special considerations involving the first or last runs.) A fraction p of the failures, approximately, should be followed by at least one success, a fraction p^2 should be followed by two or more successes, and so on. In general, therefore, there should be on average about nqp^ℓ success runs of length at least ℓ.

To gauge the likely length of the longest success run, it is reasonable to find the largest value of ℓ for which at least one run of that length would be expected to occur. We thus set $nqp^\ell = 1$ and solve for ℓ to obtain

$$\ell = \log_{1/p}(nq). \tag{1}$$

Using the closest integer to ℓ provides a simple rule of thumb for predicting the length of the longest success run in situations where the Bernoulli trials model can be applied. This formula provides a reasonable approximation as long as $nq \gg 1$.

The run lengths given by this rule of thumb are often longer than what people expect. For example, formula (1) predicts that the longest run of heads in 200 tosses of a fair coin would have length about seven. Yet few individuals, when asked to write down a simulated sequence of 200 coin tosses, include a run as long as seven consecutive heads or tails ([**5**], [**6**]; see also [**7**]). This may go a long way toward explaining the so called "hot hand" phenomenon [**2**], in which a casual observer of a sporting contest or similar situation ascribes a long run to psychological "momentum" when it is entirely compatible with natural variation.

TABLE 1 shows what formula (1) predicts for various n in coin tossing ($p = 1/2$), in ESP card matching (five symbols, so $p = 1/5$), and for the longest run of a given face in tosses of a fair die ($p = 1/6$). Values of ℓ are rounded to the nearest integer.

TABLE 1: Predicted lengths of the longest run for coin tossing, ESP card matching and die tossing

No. of trials (n)	Head runs ($p = 1/2$)	Card Matching ($p = 1/5$)	Die tossing ($p = 1/6$)
100	6	3	2
1000	9	4	4
10,000	12	6	5
1,000,000	19	8	8

The longest runs of identical outcomes

Our model so far has involved consecutive successes in a sequence of Bernoulli trials. But in many situations, we are actually dealing with multinomial trials—that is, each

trial has several equally likely outcomes, and we are interested in the longest run of identical outcomes, regardless of which outcome is being repeated.

For example, when rolling a die we might be concerned with the longest sequence in which each roll shows the same face. In the digits of π, any sequence where the same digit or the same pattern of digits repeats many times would draw our attention. Even when there are just two outcomes, it might be that the longest run of either outcome is of interest: a tail run is as notable as a head run, and a losing streak is as newsworthy as a winning streak.

Fortunately, we can adapt our model to this situation. In the case of multinomial trials, we call an outcome a "success" if it repeats the previous outcome. Now a sequence of ℓ "successes" is the same as a sequence of $\ell + 1$ consecutive identical values.

Consider a card-drawing experiment, where we treat the possible outcomes as $2, 3, 4, \ldots, 9, 10, J, Q, K, A$. An experiment might proceed as follows:

Card drawn:	J	J	8	A	K	4	2	9	8	7	7	7	3	A	4
Repeat?	-	Y	N	N	N	N	N	N	N	N	Y	Y	N	N	N

Note that the string of two repeats corresponds to a run of three 7's. Since the sequences of Y's and N's represent realizations of the Bernoulli trials model with the successes represented by the Y's, we can still apply formula (1), but need to add 1 to predict the length of the longest run of outcomes in the original sequence.

For the rest of the paper we will return to the success-run model, but now we know that we can translate when needed.

Predictability of the length of the longest run

A prediction such as those in TABLE 1 is just a "best guess" as to the likely length L_n of the longest success run in n trials. But how accurate is it likely to be?

To answer this question it is necessary to know something about the probability distribution of L_n. Here's an outline of how the approximate distribution of L_n can be found. The number of successes at the beginning of the sequence (if any) and the numbers of successes between any failure and the next one are independent, identically distributed random variables. Each of these variables has a *geometric* distribution, which means that the probability that the run has length k is given by

$$p(k) = p^k q$$

for $k = 0, 1, 2, 3, \ldots$. So the distribution of L_n is approximately that of the maximum of $\lfloor nq \rfloor$ independent geometric random variables, where $\lfloor \quad \rfloor$ represents the floor, or greatest integer function.

Due to their discreteness, the limiting distribution of the maximum of $\lfloor nq \rfloor$ independent geometric random variables with parameter p does not exist (see [**3**] and [**6**]). We can obtain an approximate limiting distribution, however, by studying the limiting distribution of the maximum of $\lfloor nq \rfloor$ independent exponential random variables, as exponential random variables are the continuous analog of geometric random variables.

To make the connection explicit, let X be an exponential random variable with parameter $\lambda = -\ln p$. The density and cumulative distribution function (cdf) for an exponential random variable are given by

$$f(x) = \lambda e^{-\lambda x} = \lambda p^x$$

and

$$F(x) = \Pr(X \le x) = 1 - e^{-\lambda x} = 1 - p^x$$

for $x \ge 0$. Now let $Y = \lfloor X \rfloor$; then

$$P(Y = k) = \int_k^{k+1} \lambda e^{-\lambda x}\, dx = e^{-\lambda k}(1 - e^{-\lambda}) = p^k q$$

for $k = 0, 1, 2, \ldots$, which is precisely the geometric probability mass function given above. Thus the distribution of the length of the longest success run can be approximated by the maximum of $\lfloor nq \rfloor$ independent exponential random variables.

Now let $X_1, X_2, X_3 \ldots$ be independent exponential random variables each having parameter $\lambda = -\ln p$, and let $M_n = \max(X_1, X_2, \ldots, X_{\lfloor nq \rfloor})$. We would like to understand the distribution of M_n. Since the value of M_n tends to increase as n increases, it will be more productive to focus on the difference between M_n and our rule-of-thumb estimate (1),

$$E_n = M_n - \log_{1/p}(nq).$$

Write F_E for the limiting cumulative distribution function of the variables E_n. We have

$$\begin{aligned}
F_E(x) &= \lim_{n\to\infty} P\left(E_n \le x\right) \\
&= \lim_{n\to\infty} P\left(M_n - \log_p(1/nq) \le x\right) \\
&= \lim_{n\to\infty} \left[P\left(X_1 \le \log_p(1/nq) + x\right)\right]^{\lfloor nq \rfloor} \\
&= \lim_{n\to\infty} \left[1 - p^{\log_p(1/nq)+x}\right]^{\lfloor nq \rfloor} \\
&= \lim_{n\to\infty} \left[1 - \frac{1}{nq} p^x\right]^{\lfloor nq \rfloor} \\
&= e^{-p^x}. \qquad (2)
\end{aligned}$$

Thus E_n, which approximates the estimation error of our rule-of-thumb estimate, has the limiting distribution given by (2). A random variable whose cumulative distribution has this form is said to have an *extreme value distribution.* Such distributions have been well studied, as they arise in the investigation of the distribution of the maximum of a sequence of independent, identically distributed random variables. The particular type of the extreme value distribution that applies here is also known as the *Gumbel distribution.* Its general form is given below.

Cumulative distribution function:

$$F(x) = P(X \le x) = e^{-e^{-(x-\mu)/\sigma}} \qquad \text{for } -\infty < x < \infty;$$

Density function:

$$\frac{1}{\sigma} e^{(x-\mu)/\sigma} e^{-e^{-(x-\mu)/\sigma}} \qquad \text{for } -\infty < x < \infty;$$

where $-\infty < \mu < \infty$ is a location parameter and $\sigma > 0$ is a scale parameter. See [**1**] for further information regarding extreme value distributions.

For the distribution that approximates the normalized maximum of $\lfloor nq \rfloor$ independent geometric random variables with parameter p, the parameters of the limiting distribution are $\mu = 0$ and $\sigma = 1/\ln(1/p)$ respectively, and the associated density is

$$f(x) = \frac{d}{dx}\left(e^{-p^x}\right) = \ln(1/p)p^x e^{-p^x} \text{ for } -\infty < x < \infty.$$

The mean and standard deviation of the distribution are

$$\begin{aligned} &\text{Mean:} \quad \gamma \ln(1/p), \text{ where } \gamma \text{ is the Euler-Mascheroni constant } 0.577215\ldots; \\ &\text{Standard deviation:} \quad \frac{\pi}{\sqrt{6}\ln(1/p)}; \end{aligned} \tag{3}$$

and the mode is zero. The density for the case $p = 0.5$ is shown in FIGURE 1 below.

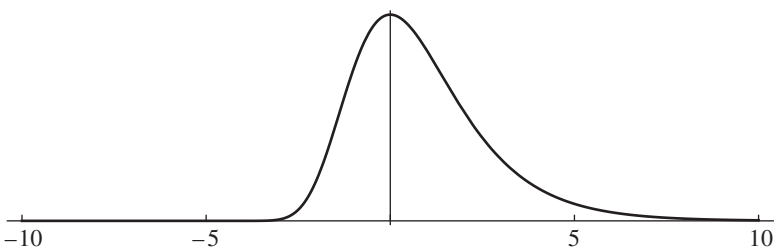

Figure 1 Exteme value (Gumbel) distribution for the case $p = 0.5$ ($\mu = 0$, $\sigma = 1/\ln(2)$)

We have established that the distribution of $L_n - \log_{1/p}(nq)$, the length of the longest success run in n independent Bernoulli trials minus its predicted value given by (1), is well approximated by $\lfloor X \rfloor$, where X is a random variable having the extreme value distribution given in (2). Therefore

$$P(L_n = \ell) \approx P(\ell - \log_{1/p}(nq) \le X < \ell + 1 - \log_{1/p}(nq)).$$

There are some important and interesting consequences of this result. First, notice that—remarkably—the approximate distribution of the prediction error, $L_n - \log_{1/p}(nq)$, does not depend on n. That means that one can predict the length of the longest success run found in, say, a million trials as well as in a hundred! Second, the standard deviation formula (3) shows that the spread of the distribution is generally quite small. This implies great predictability in the length of the longest success run.

For the case of coin tossing, for example, the length of the longest head run will very likely fall within about three of the predicted value $\ell = \log_2(n/2)$—no matter how many tosses are made. For $n = 100$, $\ell = \log_2 50 = 5.64$, so the longest run of heads (or tails) in one hundred tosses is very likely to be in the range 6 ± 3. See FIGURE 2. TABLE 2 illustrates this phenomenon for three values of n that are close to powers of 2.

In each instance there is approximately a 95% probability that the listed interval will capture the actual value of L_n. Again, what is most striking is that the prediction intervals do not get wider as n increases.

Using our theoretical results to check whether famous events such as Joe DiMaggio's hitting streak and the run of 26 spins landing on the black at Monte Carlo are beyond what should be expected presents some challenges. First of all, it is not possible to ascertain with much accuracy the value of n that should apply. Fortunately, since the predicted length of the longest run grows only logarithmically, predictions

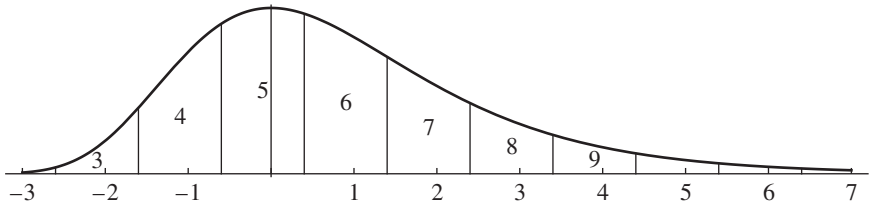

Figure 2 Approximating extreme value distribution for the longest success run in $n = 100$ tosses of a fair coin. The approximate probabilities $P(L = \ell)$ are shown graphically as areas under the curve for $\ell = 3, \ldots, 9$. Adding any positive integer i to each of these values of ℓ gives the graph for the case $n = 2^i \times 100$.

TABLE 2: Prediction intervals for the length of the longest run for coin tossing

No. of trials (n)	$\ell = \log_{1/p}(nq)$	Approximate 95% Prediction Interval
$1000 \approx 2^{10}$	8.97	9 ± 3
$1{,}000{,}000 \approx 2^{20}$	18.93	19 ± 3
$1{,}000{,}000{,}000 \approx 2^{30}$	28.90	29 ± 3

do not change greatly even when n changes by a lot. The Bernoulli trials model is not particularly suitable for baseball, given the great variation in the abilities of hitters and pitchers. Therefore predicting the likely length of the longest hitting streak goes beyond the scope of this paper. We can make a stab at investigating runs of one color in roulette, however.

A minor complication is that $p = P(\text{black})$ is different in American roulette and in the European version, being $18/38$ in the former and $18/37$ in the latter. TABLE 3 below shows the predictions formula (1) gives for both values of p and for various n, along with prediction intervals that take into account the spread of the approximating extreme value distributions around the value given by (1). These distributions have very similar spread to that shown in FIGURE 2 since both values of p are close to 0.5, therefore adding and subtracting three from the value given by (1) again gives a high probability (approximately 95%) of capturing the actual longest run length. The predictions (rounded to the nearest integer) shown in TABLE 3 are for the longest run of either black or red, which as shown earlier are obtained by simply adding 1 to the prediction for black alone.

TABLE 3: Predicted lengths and prediction intervals for the longest run of black or red in roulette

No. of spins (n)	American Roulette ($p = 18/38$)	European Roulette ($p = 18/37$)
100,000,000	25 (22, 28)	26 (23, 29)
500,000,000	27 (24, 30)	28 (25, 31)
1,000,000,000	28 (25, 31)	29 (26, 32)

The results are quite compatible with the actual run of 26; certainly that run was not suspiciously long when seen in the context of a century of experience. It seems

reasonable that the actual number of plays of casino roulette that have ever been made is likely to be between 10^8 and 10^9, but even for n well above or below this range an observed longest run of 26 would not be very surprising.

Extremely predictable cases

The smaller the value of p, the more concentrated the approximating extreme value density becomes, and thus the length of the longest success run becomes even more predictable. For example in dice tossing, the probability that the length of the longest run of any particular face ($p = 1/6$) will be one of only three consecutive values oscillates between 92.5% and 97% as n varies (except for very small n, where the chance is even higher). See FIGURES 3 and 4.

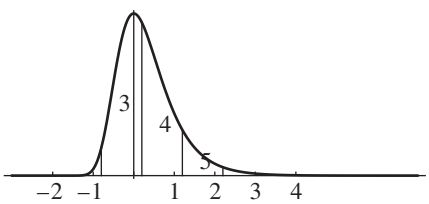

Figure 3 Approximating extreme value distribution for the longest success run in $n = 1000$ tosses of a fair die. There is approximately a 96% chance that the length of the longest run will be one of the three values 3, 4 and 5.

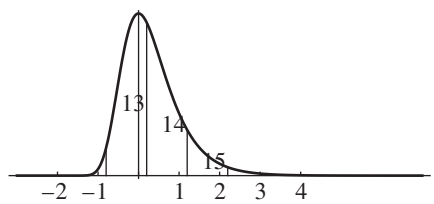

Figure 4 Approximating Extreme Value Distribution for the Longest Success Run in $n = 6^{10} \times 1000$ Tosses of a Fair Die. There is approximately a 96% chance that the length of the longest run will be one of the three values 13, 14 and 15.

In other situations where p is even smaller, we can often almost guarantee the exact length of the longest run. To see this, consider the following argument: The probability of a success run of length at least l immediately following a particular failure is p^ℓ. Let N_ℓ be the number of success runs of length ℓ or longer in n Bernoulli trials. If we condition on the number of failures n_F in the sequence (treat it as given), we obtain

$$\begin{aligned} P(L_n = \ell) &= P(N_\ell > 0, N_{\ell+1} = 0 | n_F) \\ &= P(N_{\ell+1} = 0 | n_F) - P(N_\ell = 0 | n_F) \\ &= (1 - p^{\ell+1})^{n_F} - (1 - p^\ell)^{n_F} \approx e^{-n_F p^{\ell+1}} - e^{-n_F p^\ell}. \end{aligned}$$

Writing $x = e^{-n_F p^\ell}$, we have $P(L_n = \ell) \approx g(x) = x^p - x$. Simple calculus shows that $g(x)$ is maximized at $x = p^{1/q}$. By the Law of Large Numbers, if n is large we cannot go far wrong by treating n_F as equal to its expected value, nq. Solving $e^{-nqp^\ell} = p^{1/q}$ for n yields

$$n = \frac{\ln(1/p)}{q^2 p^\ell}; \tag{4}$$

which gives

$$P(L_n = \ell) \approx p^{p/q} - p^{1/q}. \tag{5}$$

Formula (4) tells us the "special" values of n (when rounded to the nearest integer) that maximize $P(L_n = \ell)$ for a given success probability p and run length ℓ. The limit

of the expression in (5) as p approaches zero (a good calculus exercise!) is 1. Thus if p is small enough and the number of trials is somewhere near the value of n given by (4), then $P(L_n = \ell)$ is close to one. In other words, the longest success run is very likely to have exactly length ℓ.

Here is an application: Twice a day the California Lottery Commission runs a game called the "Daily 3", in which a player chooses any three digit combination. When the drawing is held, the player wins if his or her three numbers match those drawn. The player may buy a ticket that requires that the numbers be matched in order (a "straight"), or one that allows the matches in any order (a "box"). Many other states also offer this game; it has also been run illegally in many big cities for decades.

Suppose we look at the sequence of all drawings over a certain period of time for cases in which the same three number combination appears in a consecutive run of games. As shown earlier, we can compare each drawing to the previous one and define success as the case when the numbers drawn in two successive games are the same. Thus a run of length ℓ successes represents $\ell + 1$ consecutive games with the same set of numbers.

For the case of "straights", $p = (1/10)^3 = .001$. Straightforward calculations from the formulas above show that within any string of around $n = 7000$ games, the longest run of repeats will almost certainly be $\ell = 1$; that is, there will be two games in a row with the same three number sequence, but never three. The chance of this outcome (from (5)) is approximately 99.2%. This is such a high probability that if there were three or more consecutive games with the same number sequence, or no consecutive games with the same number sequence, we might have reason to suspect fraud!

The California Daily 3 has been in operation since April 1992. Checking the first 7000 plays, which cover a span of approximately fifteen years, we find that indeed as predicted, the same straight has occurred in two consecutive drawings (on six occasions), but never has the same straight occurred three times in a row.

As a second illustration, consider the longstanding question of whether the digits of π are a pseudorandom sequence, that is, can be satisfactorily modeled as a string of independent digits in which each digit is equally likely be one of the values 0 through 9. Several studies of this question have been performed over the past half-century; none have found significant evidence against the pseudorandomness hypothesis (see [**4**]). The predictability of long run lengths provides an additional and very simple test of the pseudorandomness of the digits of π.

Clever algorithms along with recent advances in high speed computing have allowed the determination of more than ten trillion digits of π, providing a very large potential database for analysis. We will consider here a much smaller number, however: the first 5,000,000 digits. This n is nearly an optimal value (found from (4) above) for predicting the run length of any two-digit pair with a high degree of certainty. Note that $p = (1/10)^2 = .01$ here if π is pseudorandom. Formula (5) indicates that if the digits of π behave as a random sequence of digits, then with approximately 94.5% probability, the longest run length of any specified two-digit pair will be three. Thus, for example, the string ... 383838 ... will almost certainly be found somewhere within the first 4,700,000 digits of π, but the string ... 38383838 ... will quite likely not be found.

Searching all two-digit pairs (facilitated by the Pi-Search Page at http://www.angio.net/pi/piquery) reveals outstanding agreement with the prediction above. For 94 of these 100 pairs the longest run is indeed three. (The other six pairs each have longest runs that are four long—the most curious being 42424242, which begins at digit 242422!) Searches of the digits of π for longest runs of patterns with other than two digits also give excellent support for the hypothesis of pseudorandomness, corroborating what other tests of π have indicated.

Values of n satisfying (4) at least approximately are special values that place nearly all of the mass of the probability distribution of the longest success run on a single value. Even for arbitrary values of n, however, this distribution lives almost entirely on at most two values when p is small. For example, similar calculations to those above show, assuming the digits of π are a pseudorandom sequence, that the chance that the number of consecutive repeats of any particular three-digit pattern within the entire currently known ten trillion digits of π is either four or five is approximately 99.994%.

Acknowledgment Thanks to Andrea Nemeth for preparation of the figures.

REFERENCES

1. Stuart Coles, *An Introduction to Statistical Modeling of Extreme Values*, Springer-Verlag, New York, (2001).
2. T. Gilovich, R. Vallone, and A. Tversky, "The hot hand in basketball: On the misperception of random sequences," *Cognitive Psychology* **17**(3) (1985) 295–314; http://dx.doi.org/10.1016/0010-0285(85)90010-6.
3. L. Gordon, M. F. Schilling, and M. S. Waterman, "An extreme value theory for longest head runs," *Zeitschrift fur Wahrscheinlichkeitstheories verwandt Gebeite* (*Probability Theory and Related Fields*) **72** (1986) 279–287; http://dx.doi.org/10.1007/BF00699107.
4. T. Jaditz, Are the digits of π an independent and identically distributed sequence? *The American Statistician* **54**(1) (2000) 12–16; http://dx.doi.org/10.2307/2685604.
5. P. Révész, Strong Theorems on Coin Tossing, *Proceedings of the International Congress of Mathematicians*, Helsinki, 1978, pp. 749–754.
6. M. F. Schilling, The longest run of heads, *The College Mathematics Journal* **21**(3) (1990) 196–207; http://dx.doi.org/10.2307/2686886.
7. W. Wagenaar, Generation of random sequences by human subjects: A critical survey of the literature, *Psychological Bulletin* **77**(1) (1972) 65–72; http://dx.doi.org/10.1037/h0032060.

Summary When data arise from a situation that can be modeled as a collection of n independent Bernoulli trials with success probability p, a simple rule of thumb predicts the approximate length that the longest run of successes will have, often with remarkable accuracy. The distribution of this longest run is well approximated by an extreme value distribution. In some cases we can practically guarantee the length that the longest run will have. Applications to coin and die tossing, roulette, state lotteries and the digits of π are given.