

The Law of the Third

Adam Atkinson (ghira@mistral.co.uk), G4G14, April 2022

The Law of the Third may be my second-favourite gambling myth. It's a distant second, since it's not remotely as good as the Samaritani Formula (described at a previous G4G), but it seems to turn up in more countries and languages, which means there's more chance G4G attendees may run into it or its victims. In French it's called "la loi du tiers", in German "Zwei-Drittel-Gesetz" (where the name has "two thirds" in it) and in Italian it's "la legge del terzo". I'd love to hear about sightings in other countries and languages.

It's a reasonably interesting example of pseudomathematics. I run into it, or am asked about it, at intervals of several years, and have been since my first encounter with the Samaritani Formula 20+ years ago. Sometime people ask me how the Law of the Third works (It doesn't!), or where the trick is, or they are incredulous that I think the Samaritani Formula is a myth and say "Next you'll be telling me the Law of the Third is a myth!". Yes, I suppose I shall!

What exactly does the Law of the Third say? Well, this isn't entirely clear since the people who spread it don't like to be, or are incapable of being, terribly clear or specific.

You will probably find two main variants of the LOTT.

(i) If you draw one item from n , n times, $n/3$ items will fail to appear.

(ii) If you draw one item from n , n times, $n/3$ items will fail to appear, $n/3$ will appear once, $n/3$ will appear twice.

If someone says one of these in exactly these terms, they are implicitly saying that draws are not independent. So your die / roulette wheel etc. must contain sensors, memory, gyroscopes, motors. Nanotechnology or magic may be involved.

And it will be claimed that this either is a method for making money playing roulette (or lotteries, or similar), or that it is a mathematical principle which allows one to find a method for etc. etc. To be fair, this is the bit that's a myth. Without too much effort, we can state a version of the LOTT which is "true but irrelevant" in that it's a true statement that does NOT allow us to win money playing roulette. In much the same way that the Samaritani Formula can be expressed in a way which is true but irrelevant.

" $n/3$ items will" may be stated as exactly $n/3$, or at most, at least, approximately, on average, ... "at most" "at least" and "approximately" are all false in any reasonable model of the universe. "often approximately" and "on average" are more or less true but irrelevant.

How do they say you can make money using this? They seem not to want to be very clear about this. Perhaps so they can sell you worthless software, books or consultancy services, or so if you can't make it work they can claim you weren't using it properly. The details don't really matter, since it can't possibly work (in any plausible model of the universe). It's a bit like someone saying they've found two odd integers whose sum is odd. No, they have not. There's no point checking particular examples: the

whole enterprise is pointless. However, I've seen a few people say things like "Go to a roulette wheel, start observing the results whenever you like and after 24 spins, ...". Why 24 spins? Because it's about $\frac{2}{3}$ of 37, so that at least $\frac{1}{3}$ of the numbers won't have appeared and you thus can't see (some versions of) the LOTT failing? But if, say, the same number has appeared 24 times you can't get enough new numbers in the next 13 spins to drag the "unseen numbers" value down enough. They then want you to imagine that .. something something in the next 13 spins? But they don't want to explicitly say that the spins aren't independent because they know that sounds ridiculous? It's very strange. I suppose as with Nigerian email scams they are targeting people who aren't paying much attention to begin with.

It's not necessary to do any mathematics at all to realise that the LOTT is nonsense. If it were possible to reliably make money playing roulette, roulette would not exist.

The LOTT seems to be some combination of an urban legend, a mistake, a con, and even a conspiracy theory. An urban legend because some of the people who repeat it don't seem to have thought about what they are saying. A con because some of the people who repeat it seem to be trying to make money, somehow, from victims of it. A conspiracy theory because some people who talk about it claim that arrogant mathematicians / rationalists / skeptics refuse to accept it, or are suppressing it, or don't want to help poor people escape from poverty. I say "mistake" for completeness, though I find it hard to believe that anyone capable of deriving the LOTT could believe it could be useful for making money at roulette or similar.

To avoid listing lots of possibilities every time, I shall pretend here that people spreading the LOTT are in fact eugenicists, trying to drive vulnerable and gullible people to bankruptcy and suicide to improve (from the eugenicists' point of view) the gene pool. Do I think this is terribly likely? Not really, but it's more fun to imagine we are combatting comic-book villains than to imagine our "enemy" doesn't understand basic probability. Besides, some of them say things like "I've been studying probability since before you were born, Adam! How dare you!". If this were not a lie, this would mean someone had studied probability for decades without understanding what independence was. That is a little hard to credit, isn't it? As Count Rugen says, "I think that's the worst thing I've ever heard. How marvellous."

As with the Samaritani Formula, some mathematicians or debunkers say the LOTT is just nonsense, and that's a little unfair. It's entirely useless for making money playing roulette etc. but it is based on misunderstanding an actual piece of mathematics, so debunking it could be a useful exercise for students, a bit like finding the mistake in the "proofs" that $1=2$ or that right angles come in different sizes.

It's of course possible that Thirdists, or some of them, believe that roulette spins / dice throws / etc. are not independent events. That would at least be interesting, if a bit odd. But many/most of them seem to claim they do not think this. Of course, it's possible that like characters in Smullyan puzzles, they believe that they believe that spins are independent, while also believing that spins are not independent.

It's a little too easy to assume without even saying so that dice and roulette wheels and so on are perfectly fair in the sense that all results are equally likely and the spins/throws/etc. are independent.

Are real coins, dice and roulette wheels perfectly balanced? Almost certainly not. Real coins and dice are not perfectly symmetrical. The designs on the two faces of a real coin are different and this most

likely introduces some bias. Dice whose pips are concavities are clearly not perfectly balanced, and 6-sided dice may not be perfect cubes. Roulette wheels in the real world may not be perfectly symmetrical. It's even plausible that all coins of a particular design could have a similar bias to each other.

Are throws/spins/etc. of real coins, dice, roulette wheels independent? Again, almost certainly not, though here the difference between reality and idealised theory seems unlikely to be large enough to matter. If you toss the same coin billions of times, the designs on the faces probably change, smooth out, wear away, or similar. The coin itself may wear away. It's at least possible to imagine that tiny changes in the shape of the coin each time it is tossed may depend on which side ends up face down. It's harder to imagine that with real coins this is measurable or exploitable. If we made coins or dice out of shortbread or plasticine the effect could matter more.

If we really wanted to make a die whose tosses were not independent, we could make a die whose insides were like those of a snail ball. (For example <https://www.perpetuum-mobile.ch/shop/produkt/schneckenkugel/1164/>). The probabilities of results after a toss would depend on previous results, and also on the time since the last throw and the ambient temperature.

It seems pretty obvious that the LOTT cannot possibly apply to dice / roulette wheels / etc. where the probabilities can be very different and/or change. If we imagine a die made so that after the first throw all later results will be the same as the first one (something inside breaks and a weight sticks to one face?), the LOTT doesn't apply. If we imagine a die where no previous result can be repeated the LOTT does not apply. If we have a die where 1 appears 999,999 times out of a million and the other numbers have the same small probability as each other, the LOTT cannot be true.

I have seen a few people say that the LOTT is an empirical result based on observations of roulette wheels. Are they saying spins are not independent? At least implicitly, they must be.

However, it seems to be more usual to see LOTT boosters say, though not in these terms, that the LOTT applies to situations where there are n equally likely results and the trials are independent. At which point it's obvious the LOTT can't help you win.

Where does it come from?

It's pretty likely that some people reading this will have worked this out already.

Let's suppose we have a die or roulette wheel or similar which generates numbers from 1 to n . They all have a probability $1/n$ of appearing and throws/spins are independent.

If we throw our die n times, what can we say about how many numbers will not appear? Obviously a new number must appear on the first throw. That number could appear n times so we'd have $n-1$ unseen numbers. Or a different number could appear every time so we'd have 0 unseen numbers. If the LOTT is based on something real it can't be a statement that exactly $n/3$ numbers will not appear. If n is not a multiple of 3 (as 37 is not), what would that even mean? It also can't be a statement that about $n/3$ numbers will not appear as any amount from 0 to $n-1$ numbers could fail to appear.

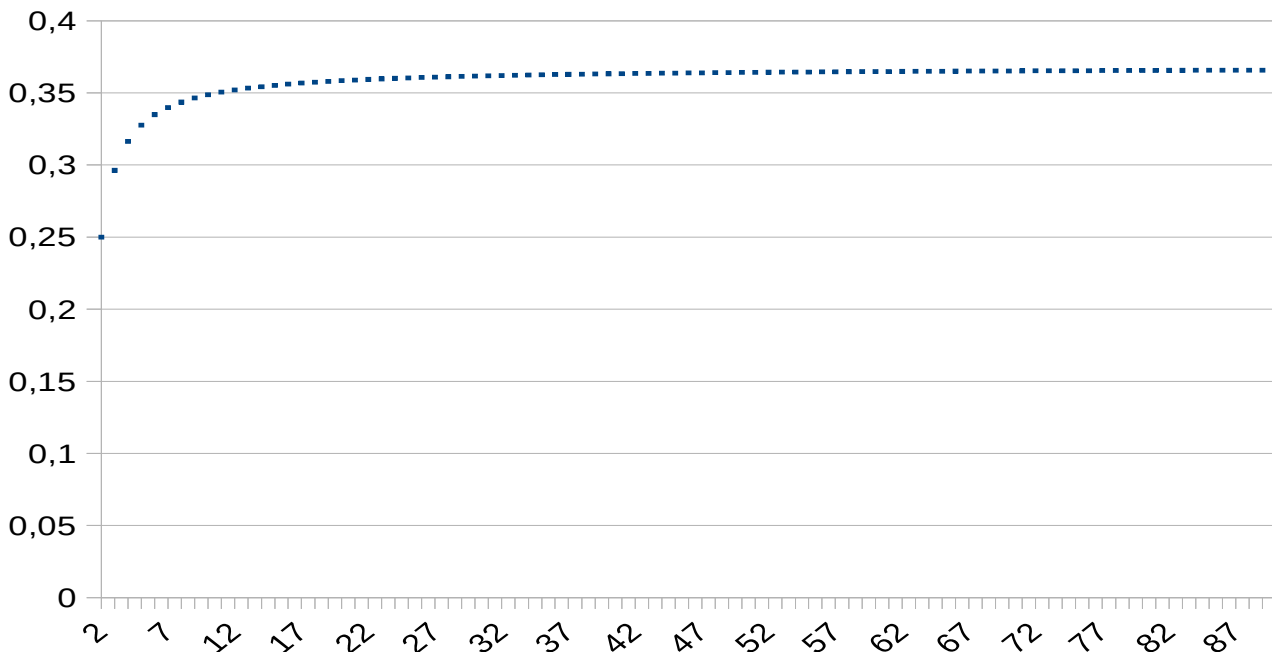
The number of unseen numbers after n throws will have a probability distribution. What can we say about it?

We can calculate its mean. Does this help us win money at roulette? If the spins are independent, of course not. However, we can engage in Mockery Through Participation. Instead of merely saying Thirdists are wrong, we can understand their “Law” better than they do. Well, better than they appear to. The ones who are con artists or eugenicists understand it perfectly and know that it’s useless, of course.

What is the probability that the number 1 does not appear after n throws? It’s $\left(\frac{n-1}{n}\right)^n$

The same holds for all the numbers. Since $E(X+Y)=E(X)+E(Y)$ whether X and Y are independent or not, the mean number of unseen values is $n\left(\frac{n-1}{n}\right)^n$. For large n , this is approximately $\frac{n}{e}$, so the name of the LOTT is wrong.

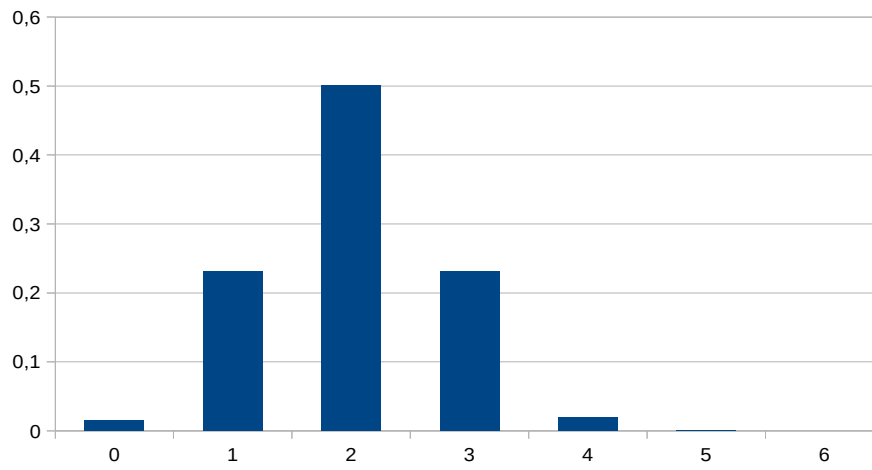
If we know n then we might as well use the exact value. Here is that value for n from 2 to 90.



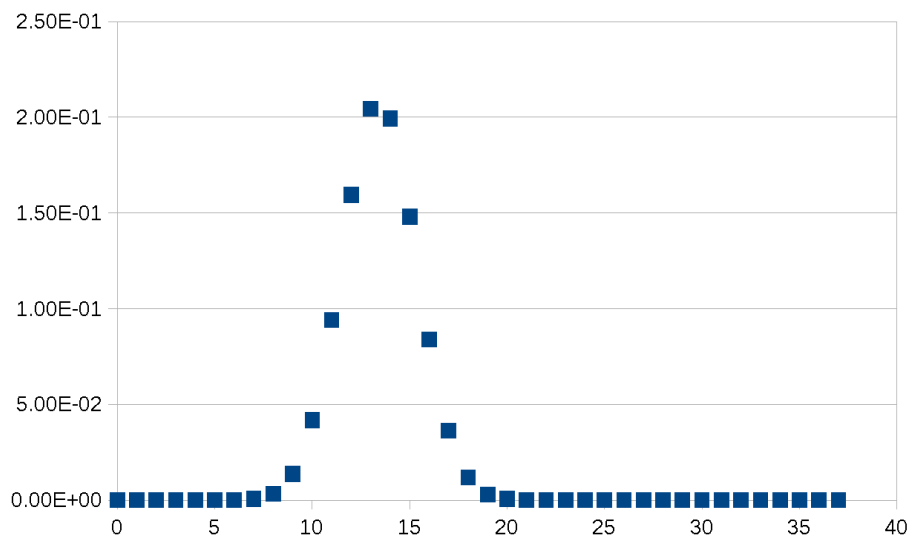
It seems clear that $1/e$ is better than $1/3$ even for quite moderate values of n . Certainly by 37 (roulette) but not 6 (a 6-sided die).

But we can calculate the entire distribution using a Markov chain. The state of the chain is the number of unseen values. We start in state n with probability 1. When we are in state i the next number could be one we have seen before (probability $(n-i)/n$) so we stay in state i . Or it could be a new number (probability i/n), moving us to state $i-1$. We turn the handle n times and look at the probabilities of the states.

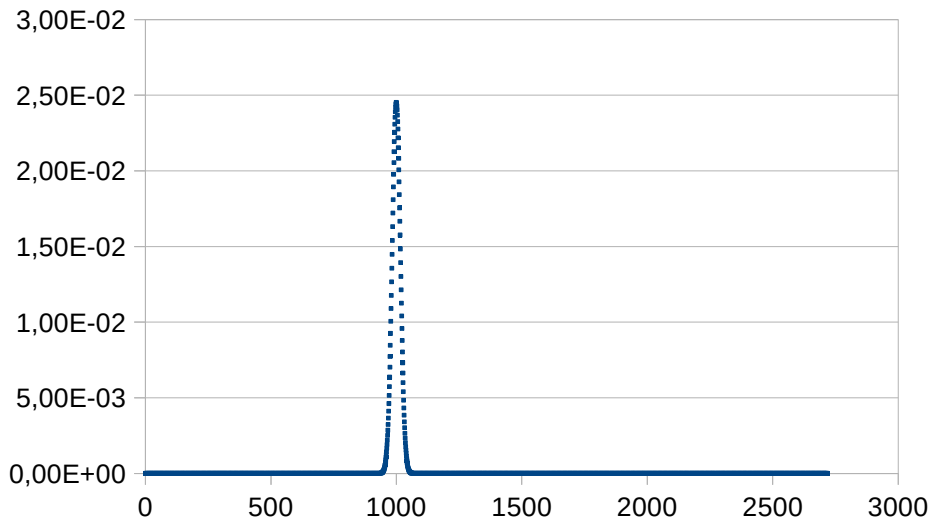
For $n=6$ we get



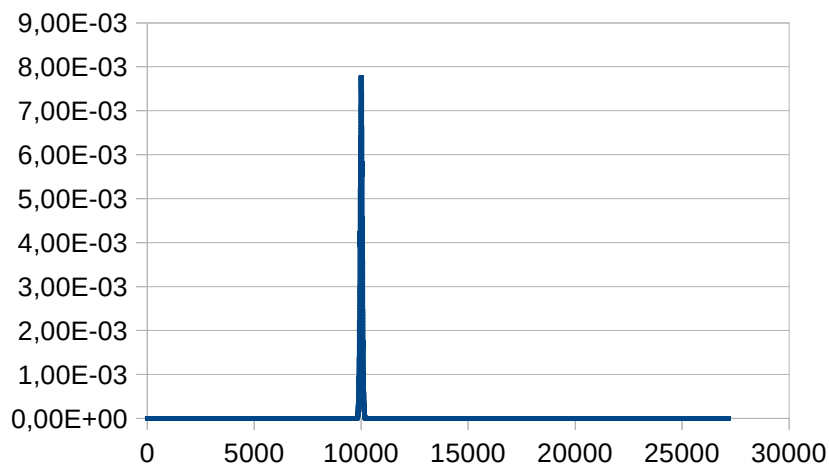
For $n=37$ (European roulette) we get



To make it pretty clear that $1/e$ is better than $1/3$ let's try $n=2718$



And, why not?, $n=27183$



As well as the mean being n/e , it looks like the mode is n/e , or the nearest integer to it, or near enough. If for some reason someone says "I'm about to throw an n -sided die n times. Guess how many numbers won't appear!" then your best chance of "winning" is to say the mode of this distribution. If you have time to calculate it exactly, great, but the nearest integer to n/e will probably be good enough if n is not too small.

It seems pretty clear that the LOTT is a misunderstanding or misapplication of this (fairly well known) result. As part of a con, it's probably better to say things which are true but irrelevant from time to time as well as falsehoods.

So some form of the LOTT, or "law of $1/e$ ", is of some use if none of the n throws have yet taken place.

If you've just watched 36 roulette draws and are asked to predict how many unseen numbers there will be in those 36 and one more draw you don't use the LOTT. It will either be the number of unseen numbers as it was after the first 36 draws or one less than that. You don't need to predict anything about the first 36 draws as you know their results.

Of course this is the "basic" LOTT.

What can we say about the extended version?

This is less fixable. Similar reasoning shows that the proportions of number appearing never, once, and at least twice are approximately $1/e$, $1/e$ and $1-2/e$. Numerically these are

0.367879441171442

0.367879441171442

0.264241117657115

And to claim that's $1/3$, $1/3$, $1/3$ seems a bit too far off.

Again if we want to look at the probabilities of various exact combinations we can model this as a Markov chain. We can call the states (i,j) where i is how many numbers have appeared once and j is how many have appeared at least twice. We start in state $(0,0)$ with probability 1. If we are in state (i,j) , which states do we go to next? With probability j/n a "double" appears again and we stay in the same state. With probability i/n a "singleton" appears again and we move to state $(i-1,j+1)$ and with probability $(n-i-j)/n$ a new number appears and we move to state $(i+1,j)$. Turn the handle n times and see what you get.

LOTT and lotteries

I've also seen people "applying" the LOTT to lotteries. Specifically, to single wheels of the Italian "lotto". Each wheel in that is a 90 choose 5 lottery. So a given number has a probability of $1/18$ of appearing in any given draw. After 18 draws you *could* get each number once. So some people say the LOTT means that after 18 draws 30 numbers will not appear. Or approximately 30. Or something. Who knows? I don't think they do.

I don't see how they would calculate the probabilities of all possible amounts of unseen numbers this way. They probably wouldn't. Maybe, implicitly, they're using a "radioactive decay" model where you have 90 particles in a box and at each step each one has a $1/18$ chance of vanishing. What is the distribution of the number of particles remaining after 18 steps? This is a pretty poor model of a 90 choose 5 lottery of course. After 1 step you could have any number of particles between 0 and 90. And indeed after 18 steps. The states are of course the numbers from 0 to 90 and the transition probabilities can be calculated using binomial distributions.

It would seem closer to reality, though still wrong, to model 18 draws of a 90 choose 5 lottery as a 90 number roulette spun 90 times. This is wrong because draws cannot contain duplicates: the first draw will contain 5 distinct numbers, so states 86 to 89 can never happen. It might be amusing to compare the

distributions obtained with these two wrong models.

Of course we can do "better" than this in the sense that we can perform more accurate and complete useless calculations, again in a spirit of Mockery Through Participation. If we have a 90 choose 5 lottery, what is the distribution of the number of unseen numbers after 18 draws?

Again, Markov Chain. The state is the number of unseen numbers. Initially it is 90 with probability 1. In state i you can go to states $i, i-1, i-2, i-3, i-4$ and $i-5$ (if $i < 5$ only some of these are possible, of course).

The probability of going to state $i-j$ is $\frac{\binom{i}{j} \binom{90-i}{5-j}}{\binom{90}{5}}$. Turn the handle 18 times and see what you get.

Multiples

I've also seen people talking about exact multiples of n draws/spins/tosses and how the fraction of unseen numbers will "be" (in whatever sense they mean) $\frac{1}{3^n}$. The difference between 3 and e will eventually make itself felt. It's a little odd that they seem to think that making exact multiples of n draws etc. has some special meaning.