47 Comments   **Mark Fickett Art**   💬 **Mark Fickett** ⌄

♡ **Recommend**   🐦 **Tweet**   f **Share**   Sort by Best ⌄

Join the discussion…

SmilodonsRetreat • 3 years ago

Have you tried any of the machine aluminum dice? If you are local, I'd be willing to let you borrow them. If not, I can get you the names and e-mails of some people who have them.

1 ∧   ∨ • Reply • Share ›

> **Mark Fickett** **Mod** → SmilodonsRetreat • 3 years ago
>
> I haven't tried any metal dice. With the paperboard bucket, and given how it's just taped to the servo motor, I don't think this setup is strong enough. If I make a v2, I may lasercut MDF or something a little tougher, so I can test some metal dice. Thanks for the offer though!
>
> ∧   ∨ • Reply • Share ›
>
> > **SmilodonsRetreat** → Mark Fickett • 3 years ago
> >
> > So I guess the solid tungsten d6 is right out then? ;)
> >
> > ∧   ∨ • Reply • Share ›

Matt Brown • 3 years ago

Is there any possibility that very slightly different accuracies in the computer measuring the roll outcome for dice of different font, background colour, translucence, die size etc may be affecting this?

1 ∧   ∨ • Reply • Share ›

> **Mark Fickett** **Mod** → Matt Brown • 3 years ago
>
> I think the categorization is accurate. When I'm running the grouping step, I do sometimes see false positives, but I generate a summary image where not-as-good matches are nearer the front of each list of matching images, so it's pretty easy to spot them. When I do see false positives, I just increase the match threshold and re-run the script.
>
> If you check out the github repository, each die's subdirectory in the data/ directory has a summary.jpg which is essentially my QA step; all the dice on one row should be the same. If you spot any mismatches let me know! Example: https://github.com /markfick...

So, different dice do require different parameters around detecting matches, but I adjusted to each one.

1 ^ | ∨ • Reply • Share ›

laszlokorte ➜ Mark Fickett • 2 months ago

Look at the 2nd row/2nd column from top left in the picture you linked

^ | ∨ • Reply • Share ›

Mark Fickett Mod ➜ laszlokorte • 2 months ago

Good catch! Comparing to the other 16s, it looks like it's actually a little bit tipped while they appear flatter, I bet that's why it was miscategorized.

^ | ∨ • Reply • Share ›

Michael Liesenfelt • 3 years ago

You are not adding adequate entropy to your system when the dice is being 'rolled' from almost the same position every trial, the normal vectors of the first impact points of every 'roll' are very similar, and the number of bounces per roll can be counted on one hand. I would recommend using more energy per 'roll' and a level resting area before each 'roll'. Your low potential area resting area/circle and OpenCV photography method are great! The entropy of your system is proportional to the product of the high potential energy rest area, low potential energy resting area, and total potential energy added to the system per roll.

1 ^ | ∨ • Reply • Share ›

Mark Fickett Mod ➜ Michael Liesenfelt • 3 years ago

I didn't find a great way to test the machine itself. I'm sure you saw the sequence analysis in the "Roller Randomness" section, but as others commented below, while it seems like you'd see a clear pattern if one was there I really didn't do much analysis in that vein. (If someone wants to play more with the sequence data and needs help getting at it, I'm glad to help.)

But I think there may be more energy in the system than you're allowing. When the servo motor flips the tub back upright, it does so fast enough that the die doesn't just roll back down, it bounces off the far side of the container, usually hitting the LEDs / wire in the process. So there's both the potential difference, and the additional kinetic energy. Regardless, could you point me to any resources about analyzing entropy in systems like this?

I'm also keeping an eye out for an opportunity to test one of these dice via a different method / machine, to compare behavior in different contexts.

1 ^ | ∨ • Reply • Share ›

Evan Verworn ➔ Mark Fickett • 2 months ago

Isn't that 16 in the top left (x:2, y:2) incorrectly categorized as a 3?

∧ | ∨ • Reply • Share ›

goulo • 3 years ago

Admirable geekiness! Very cool.

1 ∧ | ∨ • Reply • Share ›

Mark Fickett **Mod** ➔ goulo • 3 years ago

Thanks!

∧ | ∨ • Reply • Share ›

Holger Peters • 2 months ago

Have you tried making a Bayesian inference with a multinormal likelihood and a Dirichlet-prior? It would give you distributions for the probability for each side of the die, i.e. built-in certainty estimation that you could definitely use to plot an error bar.

∧ | ∨ • Reply • Share ›

Mark Fickett **Mod** ➔ Holger Peters • 2 months ago

That's way beyond my stats knowledge, unfortunately. If you wanted to try that out on the raw die-roll data (for example https://github.com/markfick..., I'd be curious what you find. I did implement using bootstrapped subsamples to generate confidence intervals (https://github.com/markfick..., but haven't updated this page with the results yet. (You can see drafts of new charts with those confidence intervals at http://www.markfickett.com/... .)

∧ | ∨ • Reply • Share ›

Ben • 3 years ago

With tweaks to the image recognition, could you roll multiple dice simultaneously? (ignore if this is discussed above -- I didn't see it and tried a couple text searches).

You couldn't roll the same type or color of dice together, unless you could somehow "mark" the die without affecting rolls. Maybe spray a little aerosol sunscreen on one and use a UV filter. But spray would introduce an unwelcome variable.

Could also segment the bucket into quadrants.

Great project!

∧ | ∨ • Reply • Share ›

Mark Fickett **Mod** ➔ Ben • 3 years ago

I haven't tried to do multiple dice at once. I like the idea of UV-marking one die; I think you can get UV LEDs. However my guess is it'd be more complication than payoff: shading on part of one die or on the background is already sometimes a problem; paint on the rolling container wore off on dice so probably UV paint would wear off onto the tub; my code for picking out one die from the bucket is very heuristic-based and probably wouldn't generalize to two especially if they ended up next to each other (UV marking notwithstanding); and the documented setup is pretty much fire-and-forget as far as running it so it wouldn't be a huge benefit to run two dice at once (cool as it'd be). Square dice (and more of a machining budget) has already brought this to fruition, though -- have you seen this?
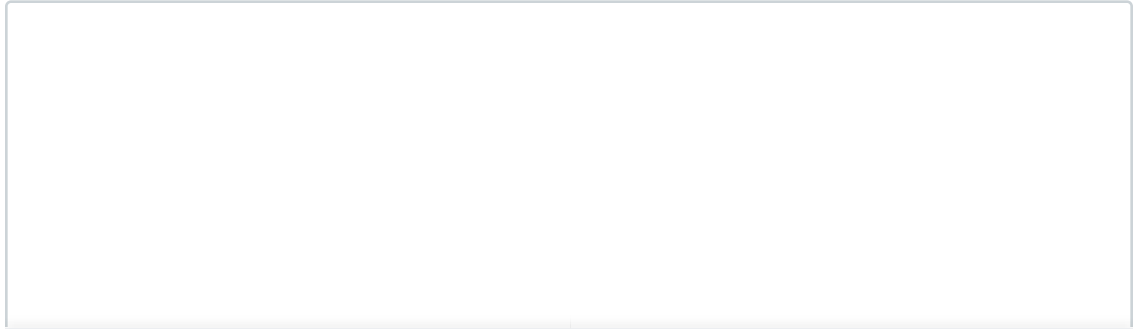
see more

∧ | ∨ • Reply • Share ›

Daniel Fisher • 3 years ago
Hey I have an idea for a test how can I get in contact with you to discuss this?

∧ | ∨ • Reply • Share ›

Watson Ladd • 3 years ago
To compare distributions against uniform, you should use a chi-squared test.

∧ | ∨ • Reply • Share ›

Mark Fickett **Mod** ➔ Watson Ladd • 3 years ago
Hi, thanks for reading. I actually looked into the Chi-squared test, and I don't think it tells me what I want. Chi-squared says "how likely is this to have come from a random source?" (as you said, comparing against uniform). That's good for answering the binary question "is this a fair die?" with some level of assurance but as you can see, basically no dice are fair. What I'd really like to know is how good a predictive model I have developed from the rolls I've recorded; and how the unfairness of different dice compare.

For example, take the first two Wiz dice (markfickett.com/dice#wizdice), from High City Books. With a Chi-squared test, they both have $p < 0.000000$ (using SciPy's library function, at https://github.com/markfick... if you're curious). But the two

distributions are actually very different: the translucent blue die has extremely low frequency of rolls for 3 and 17, whereas (while still uneven) the opaque purple die doesn't have outstanding notches or peaks. A standard deviation of the frequencies does somewhat reflect this difference (0.28 and 0.14 respectively) which is why I chose that as my overall "fairness" measure.

However I haven't found a good measure of how good a predictive model I have. People have suggested error bars on the histograms, which is what I'd like, but I haven't found a formula that applies to multiple potential outcomes (as opposed to estimating one value within a population).

︿ ∨ • Reply • Share ›

Stacey • 3 years ago

For my engineering degree, my capstone course involved designing a machine to roll dice repeatedly for this type of a thing... I'll have to see if I can find our reports on our setup - we did it a bit differently than yours and didn't have LEDs inside the rolling chamber (which could help solve the marring mentioned)

︿ ∨ • Reply • Share ›

Mark Fickett **Mod** ➜ Stacey • 3 years ago

That'd be great to see! In my initial tests I found that uneven lighting was a problem, though at that point I was matching images and not features, so I'm curious what your overall lighting setup was (in particular).

︿ ∨ • Reply • Share ›

Stacey ➜ Mark Fickett • 3 years ago

Apparently all I can find is our final presentation... we didn't include any statistics in this presentation, but it does include a simple physical layout diagram on slides 14 and 15. Ours had a "cement truck" style tumbler that would rotate to roll the dice.

https://www.dropbox.com/s/b...

︿ ∨ • Reply • Share ›

Mark Fickett **Mod** ➜ Stacey • 3 years ago

Thanks for sharing that!

︿ ∨ • Reply • Share ›

Wes • 3 years ago

Surprised the Gamescience die wore down. Can I get more details on when it was made and by whom (Gamescience or Gamestation...)

Mark Fickett Mod ↱ Wes • 3 years ago

It didn't wear down so much as mar; you can see in the picture that it has little scratch marks on the surface. My guess is that's partly because of how the LEDs stick into the rolling container: much pointier than the average gaming table.

I got both the white and the black Game Science d20 from gamesciencedice.com as individual dice (http://www.gamesciencedice....).

∧ | ∨ • Reply • Share ›

Xevioso • 3 years ago

I had another thought regarding your process...you may have just found a lucrative career for yourself. You may want to contact the Nevada Gaming Commission and show them what you have done, and offer to run some tests on their casino dice. I am not sure what methodology they use to verify their dice are fair, but it's my understanding that by law they have to make their dice fair, so they fill in the pips with plastic so that one side is not weighted more than the other. There are also no rounded edges. It may be the case that they assume the dice are fair because they have filled in the pips with plastic, but is it different plastic than the rest of the dice, and does the weight of ink on the pip make any difference in the rolls?

You'd have to refine your methodology a lot more, but I suspect they might be interested in what you are doing. You could offer to run some tests on their dice to make sure they are truly rolling randomly. If there's any statistical deviation *at all* they would be curious to know. And if they aren't, some enterprising lawyer might want to speak to you if you publish results that the dice are in fact not truly random, so that such evidence might be used in a lawsuit against a casino. Either way, your little experiment might lead to some interesting opportunities for you if you pursue it further.

∧ | ∨ • Reply • Share ›

bughunter • 3 years ago

This is awesome. I don't recall any prior art but I'd be surprised if someone hasn't performed studies like this before, even if rolling and recording numbers manually... they may give you ideas and techniques for statistics.

I do suggest you take larger samples of each company's dice to see if, for instance, Game Science has consistently fairer dice, or if Wiz Dice's variability is real.

∧ | ∨ • Reply • Share ›

Mark Fickett Mod ↱ bughunter • 3 years ago

Thanks! There's actually a fair amount of prior art (I linked a few inline in the intro blurb at the top). I've even read some offhanded comments in other discussions that hint at casinos having made machines like this for their own use, but I couldn't find a

writeup about such a thing.

As to testing more dice per manufacturer, there are two problems. One is that I'd have to buy them (anyone want to donate a bunch of dice?). The other is that I'm using a dSLR with a mechanical shutter, and the shutter's lifetime is already in jeapordy. camerashuttercount.com says I started at shutter count 11545 (12% of a D90's expected lifetime), and am now at shutter count 96087 (96%); even buying a used dSLR for a couple hundred dollars, if the lifetime estimate is accurate it's much more expensive to wear out the camera than to buy the dice. I have a Raspberry Pi with a camera module, though, so I may port the system over if I continue testing.

My hope is that by documenting the machine (really pretty simple construction) and releasing the code, I'll make it easier for someone else to take it the next step!

⌃ | ⌄ • Reply • Share ›

Owen Ozier • 3 years ago

Very cool. I teach a statistics example just like the one you describe. Your sample size is large enough that asymptotics should apply. An analysis I would do is to ask, for each die and for each number on that die, is the frequency of that number statistically distinguishable from 0.05? Drop me a line to discuss offline.

⌃ | ⌄ • Reply • Share ›

David • 3 years ago

I'd be willing to purchase that white game science die from you ;-)

⌃ | ⌄ • Reply • Share ›

Mark Fickett **Mod** ➜ David • 3 years ago

I've listed it on eBay! http://www.ebay.com/itm/-/1...

⌃ | ⌄ • Reply • Share ›

Aaron • 3 years ago

Would a hard surface vs paper cup bottom impact the randomness of the die rolls?

⌃ | ⌄ • Reply • Share ›

bughunter ➜ Aaron • 3 years ago

I would recommend he NOT use a hard surface if he's rolling dice over 8000 times each. It'll wear off the corners and vertices and the dice and possibly introduce a drift in results over time. I do suggest a more solid cup, but lined with felt.

In the 70's, when I first played RPGs, dice were made with less durable plastic. We avoided rolling on tables and other hard surfaces, preferring to roll on the battle mat

or books, and we replaced our dice every so often. Eventually someone introduced a "permanent" d20, made of harder stuff, and the rumor was that the "permanent" plastic was there all along, but dice makers wanted to sell razor blades, not razors.

∧ | ∨ • Reply • Share ›

Tom Mornini ➦ bughunter • 3 years ago

Dice wear is a fact of life. I want to know how dice perform over their life, wear and tear included!

∧ | ∨ • Reply • Share ›
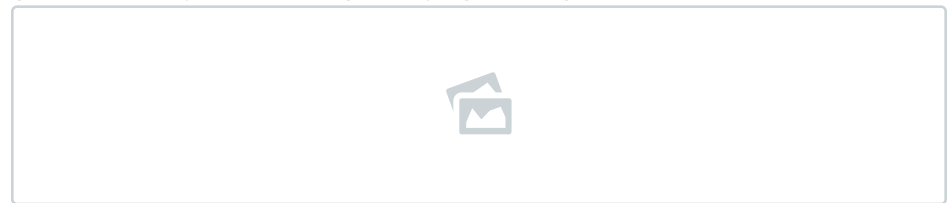
Mark Fickett **Mod** ➦ Tom Mornini • 3 years ago

Even with the dice bouncing off of the wires/LEDs inside the bucket, the only one that showed obvious *visible* wear was the black GameScience d20. (And it showed wear after only a dozen or so rolls.)

However, since I recorded not just counts but sequences, it's pretty easy to re-evaluate only a portion of the rolls.

Attached is a graph comparing the first half of the opaque purple Wiz Dice d20's rolls with the second half. The two halves of the recorded sequence differ fairly significantly at 3, 6, and 12. That probably means if there was some visible wear, it was near their opposites: 18, 15, and 9. But I don't see any obvious differences in images of those sides from near the beginning versus near the end of the roll sequence.

A different piece of evidence, though, is that the bucket did have a scattering of colorful dust in the crevices by the end of rolling 29 dice. (Unfortunately that didn't photograph well.)

∧ | ∨ • Reply • Share ›

Tom Mornini ➦ Mark Fickett • 3 years ago

Thanks!

I didn't intend to give you an assignment, was just pointing out the benefit of hard surfaces.

I suppose in casinos dice are rolled on felt for a reason. :-)

∧ | ∨ • Reply • Share ›

Jason M • 3 years ago

Extremely well done, Mark! I'll admit that I've been brainstorming basically the exact same setup myself for the past couple of years (or very, very similar at any rate). However, my idea never got out of my head, whereas you actually executed. Kudos to you!

∧ | ∨ • Reply • Share ›

Mark Fickett **Mod** ➔ Jason M • 3 years ago

Thanks!

Just as I was posting, I came across Timothy Weber's work, which is also startlingly similar: timothyweber.org/dieroller . However he hasn't posted his results yet (and says his machine is currently out of comission), so we'll have to wait to compare notes.

∧ | ∨ • Reply • Share ›

Xevioso • 3 years ago

This is truly fascinating. However, I have a question for you. I'm wondering if you can extend your tests further to a specific type of d20 die. The game "Magic the Gathering" often makes use of d20 "spindown" dice, which are dice with the numbers going sequentially from 20 to 1. This is because these dice are used as life counters in the game (in addition for rolling for random things), and it's easier to change your life from, say, 20 to 17 if you only have to slightly adjust the die, rather than continually looking for the number on the die because it's placed (seemingly) randomly.

I once encountered a person who said he did not want to roll a spindown die to determine who would go first in the game, because he thought they were unfair. His logic, which I thought ridiculous at the time, was that dice with a greater concentration of double digit numbers on the top of the dice (i.e, 16-17-18-19-20) would be unfair, because slightly more plastic from the die had been removed to deboss (rather than emboss) those numbers onto the die, as opposed to the amount of material that had been removed to deboss the single digit numbers on the opposite side of the die (1-2-3-4-5). Thus, because there was slightly more plastic on the side of the die with single digits, the die would be ever so slightly weighted in favor of higher numbers appearing on the top. As a result, this person didn't want to roll a "random" die against a "spindown" die.

This seems somewhat ludicrous to me, because even if it's true it's probably such a slight difference as to have virtually no effect in most rolls. However, the fact that you have this machine set up means you can test this for the entire MTG community. Please do so, for SCIENCE! Go down to your local comic book store, and ask them if they have any Magic

the Gathering spindown dice available. Then, please test away, and report back! People on this website would be eager to know: http://www.mtgsalvation.com/

For extra credit, you may note that the MTG spindown dice don't have a 20 on top. They have the symbol of the set from when the dice was printed; currently it's "Rise of the Eldrazi" but the spindown dice have had symbols printed on them going back ten years I think. So if you are able to get ahold of multiple spindown dice from different sets and if you can test them, the entire MTG community will know specifically which spindown die from which set is the most unfair! That is, if any of them are unfair. Thanks!

⌃ | ⌄ • Reply • Share ›

Mark Fickett **Mod** ➔ Xevioso • 3 years ago
I would like to test some spindown dice! After buying two 7-die sets, and three grab-bags of d20s, I decided I had to draw the line somewhere. But if someone wants to send me some MTG spindown dice I'd be glad to run them through the machine.

⌃ | ⌄ • Reply • Share ›

Xevioso ➔ Mark Fickett • 3 years ago
ok...what is your address

⌃ | ⌄ • Reply • Share ›

Mark Fickett **Mod** ➔ Xevioso • 3 years ago
If you don't mind sending me an e-mail at mfseas@gmail.com I'll send back my address. (And thanks!) Disqus doesn't seem to have a PM feature, and I don't want to post my mailing address publicly.

⌃ | ⌄ • Reply • Share ›

Lee LaFond ➔ Mark Fickett • a year ago
I would love to see the results on this as well. Did you ever take a look at spindown dice? One alternative would be to re-map the die results (normal d20 arrangement to spindown) from your original study and re-run the statistics to see if it makes a difference. Thanks!

⌃ | ⌄ • Reply • Share ›

Doru Constantin • 3 years ago
Very nice work! I am not sure what method of analysis would be the most appropriate, but the sequence matrix deserves a closer look. You mention the possibility of one side being preferentially followed by another one, but you only investigate the simple case of the two sides being equal (preference for the diagonal). However, the correlation could be less obvious, such as one side being followed by its opposite, or by an adjacent one.

⌃ | ⌄ • Reply • Share ›

Jason M ➔ Doru Constantin • 3 years ago

You're right, but I'm nearly positive that any such design flaw over the course of thousands of rolls would either disappear completely or stick out like a sore thumb.

∧ | ∨ • Reply • Share ›

Mark Fickett Mod ➔ Jason M • 3 years ago

I agree with both of you; I'd like to see a better analysis of the sequence (and glad to help someone grab that data), but it seems like if there's a bias in the roller it would show up more clearly.

∧ | ∨ • Reply • Share ›

Doru Constantin ➔ Mark Fickett • 3 years ago

With enough data, one could build the sequence matrix over N rolls and repeat the process M times. Then compare the positions of the maxima and minima in the M matrices. Otherwise, at least a chi-square test line-by-line (once side k is up, are the subsequent throws uniformly distributed?) would be nice. For an example, see:

https://en.wikipedia.org/wi...

∧ | ∨ • Reply • Share ›

✉ Subscribe   Ɗ Add Disqus to your siteAdd DisqusAdd   🔒 Disqus' Privacy PolicyPrivacy PolicyPrivacy