A Possible Alternative Model of Probability Theory?

1 Introduction:

We have alternative models of geometry (non-euclidean, projective, etc), alternative models of set theory (with axiom of choice, without), alternative models of calculus (non-standard analysis) – so why not alternative models of probability theory?

By adopting an alternative type of infinitesimal calculus (which I call "***dx-less***") it seems that an alternative model of probability can be constructed. Generally it seems to be the same as the standard model. But for a certain class of distributions, small deviations from the Central Limit Theorem appear.

If such alternative models prove acceptable, then the question of what models operate in what physical situations will become an empirical exercise.

2. Description of dx-less calculus

Consider the alternative type of infinitesimal calculus I call "*dx-less*". It is exactly like standard infinitesimal calculus except:

i) you toss the *dx* bit at the end away (hence the name "*dx-less*") and
ii) you evaluate f() values at the midpoint of all subintervals.

Strangely, you get convergence for many functions over certain intervals.

For the moment, lets just concentrate on the interval (0,1).

Then

$$\int_0^1 f(x) \overset{def}{=} \lim \{f(1/2), f(1/4)+f(3/4), f(1/6)+f(3/6)+f(5/6),...\}$$

(note: no *dx* at the end)

For instance,

$$\int_0^1 (1+\ln(x)) = \lim \{1+\ln(1/2), 1+\ln(1/4)+1+\ln(3/4),...\} = \ln(\sqrt{2})$$

Easy!

Playing around with various functions (whose standard integrals over interval(0,1) equal zero, obtained from a table of integrals) you'll find that logarithmic functions often converge over the interval(0,1) - but not always. Non-logarithmic functions in general are much less likely to converge. Go check.

But notice: sometimes the value of convergence differs from that of the standard integral. For instance:

$$\int_0^1 (1+\ln(x))\mathrm{dx} \;=\; 0 \text{ but } \neq \int_0^1 (1+\ln(x)) = \;\ln(\sqrt{2})$$

From such a difference, new probability theories might be made.

Now consider the problem of wanting to determine the median of a very large sample (say a billion) of 1+ln(ran#) where ran(#) is a random number between 0 and 1.

Under the Central Limit Theorem, the answer should be 0. But with dx-less calculus the answer should be closer to ln(sqr(2))=0.344... . A small difference, but a difference none the less.

To construct a new probability theory, we use functions of random numbers between 0 and 1 – written f(ran#) or f(r) for short – rather than standard probability density functions – namely pr(x).

For instance, the random function (ran#)^2 is equivalent to the probability density function pr(x)=1/2sqr(x) for 0<x<1.

In general you switch from f(ran#) to pr(x) by:

1. Replacing ran# with x giving f(x)
2. Inverting f(x) to give the cumulative density function cpr(x)
3. Differentiating cpr(x) to give pr(x)

$$f(ran\#) \xrightarrow{replace} f(x) \xrightarrow{invert} cpr(x) = f^{-1}(x) \xrightarrow{d/dx} pr(x) = \frac{d}{dx} f^{-1}(x)$$

To go from pr(x) to f(ran#):

1. Integrate pr(x) to give cpr(x)
2. Invert cpr(x) to give f(x)
3. Replace x with ran# to give f(ran#)

So now we have something to play with. If you have a pr(x) you have a f(ran#) and vice-versa.

In general, pr(x) maps the sample space onto the real number line while f(r) maps the interval (0,1) to the sample space.

Compare some results of the new model of PT with those of the old:

	<u>Old PT (pr(x) based)</u>	**<u>New PT (f(ran#) based)</u>**
E(x)	$= \int_{-\infty}^{\infty} x * pr(x)dx$	$= \int_{0}^{1} f(x)dx \ (not \ = \int_{-\infty}^{\infty} x * f(x)dx)$
Constraints on pr(x), f(ran#)	$i)\ pr(x) \geq 0 \ for \ all \ x$ $ii) \int_{-\infty}^{\infty} pr(x)dx = 1$	- virtually none? - f(x) just has to be a real function on(0,1) - can be as "wild" as you like?

For example:

$$E(ran\#^2) = \int_0^1 x^2 dx = \frac{1}{3}x^3\Big|_0^1 = 1/3$$

$$and$$

$$E(\frac{1}{2} * x^{-1/2}, 0 < x < 1) = \int_0^1 x * (1/2) * x^{-1/2} dx = \int_0^1 (1/2) * x^{1/2} dx = (1/3)x^3\Big|_0^1 = 1/3$$

In general, another way to calculate E(g(x):f(ran#)) is as follows:

$$E(g(x) : f(ran(\#)) = \int \frac{g(x)dx}{f'(f^{-1}(x))}$$

<u>For example</u>:

$$let \ f(ran\#) = (ran\#)^2$$
$$then$$
$$f(x) = x^2 \ (for \ x = 0 \ to \ 1) \Rightarrow cpr(x) = f^{-1}(x) = \sqrt{x} \ (for \ x = 0 \ to \ 1)$$
$$and \ f'(x) = 2x$$
$$thus$$
$$E(x) = \int \frac{x\,dx}{f'(f^{-1}(x))} = \int \frac{x\,dx}{f'(\sqrt{x})} = \int_0^1 \frac{\sqrt{x}}{2} dx = 1/3$$

So far , so good.

But differences seem to arise for the Law of Large Numbers (LLN) and Central Limit Theorem (CLT).

Law of Large Numbers

Old Prob Theory	New Prob Theory
$for\ S_n = x_1 + x_2 + \ldots + x_n$ $(x_i = \text{ random variables})$ $S_n \to n\mu \text{ as } n \to \infty$	$\sum_1^n f(ran\#_i) \to \int_0^1 f(x)\ as\ n \to \infty$ $for\ suitable\ f(x)$

Central Limit Theory

Old Prob Theory	New Prob Theory
$\Pr\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} < \beta\right) \to \Phi(\beta)\ as\ n \to \infty$	$\Pr\left(\frac{\sum_1^n f(ran\#_i) - \int_0^1 f(x)}{\sigma\sqrt{n}} < \beta\right) \to \Phi(\beta)?\ as\ n \to \infty$

Now here's the problem: For certain distributions, the normal curves are slightly shifted when using the new technique compared to the old. For example, f(r)=1+ln(ran#) has the normal curve centered on ln(sqr(2))=0.344... not 0 as per standard probability theory.

For many months I tried to add extra terms to either the old expression or the new to make them equivalent. But something didn't feel right. I was having to decide which one of them was "right" and which was "wrong". But I couldn't figure out which was which.

I tried simulations. I tried solving hideous sequences of ever growing multiple integrals. In the end I gave up. Nothing was providing me a definitive answer.

And then one day I had the heretical thought: what if they were ***both*** right?

Actually, they didn't have to be "right". They just had to be ***consistent***. Or at least not ***outrageously inconsistent***. Cantor had worked that out. There is no right or wrong in mathematics – just consistent, inconsistent, or of unknown consistency.

And as far as I can make out there's no great inconsistency I can find in either (although who knows what the future might bring).

So... do we have 2 different models of probability here with slightly different results regarding the Law of Large Numbers and Central Limit Theorem?

Your thoughts, please.

If we do, then the question of applicability in various physical phenomena arises.

And: could there be other models? If so, how many? With what properties?

If you can generate alternative non-euclidean geometries by adopting an alternative metric, why shouldn't you generate alternative probability theories by using a non-standard ("dx-less" or similar) calculus? And, in theory, there may be an infinite number of such alternative calculi. So how many of those might be used to construct a probability theory?

Please feel free to investigate this matter.

D Williams (Nov 2016)
Coonabarabran NSW 2357
everythingflows@hotmail.com