Comments on "Tossing a coin"

Your "abstract" isn't an abstract. Suggestion: something like this:

"A sequence of tosses of an unfair coin determines the bits in a binary expansion. We explore the some of the remarkable features of the cumulative probability distribution of this random variable."

Can you see the difference? A terse and partial description of the object of study and of the results.

In the first paragraph, I think you should introduce the parameter p rather early, and certainly before it occurs as a subscript. I think you should also point out the special cases $p = 0$, $p = 1$, and $p = 1/2$. Also we might hear that (as long as p is not 0 or 1) the ambiguity of the binary representation of dyadic numbers doesn't affect the value of f_p .

Computation of $f(1/3)$ doesn't really belong in the introduction. After you have defined your function (paragraph 1), I think you should refer to the graphs. You could put values of p both $> 1/2$ and $< 1/2$ on the same graph (say which order the graphs occur in).

Then state your basic results, precisely, but without proofs. [I'd love to see a proof that the function is not differentiable at dyadics, and that if it is differentiable at a then $f'(a)=0$ (even for the special case $a = 1/3$). From what is written, f_p could be differentiable everywhere.]

Then, in the next section, collect some basic properties of f_p : monotonicity, functional equations, relation between f_p and f_{1-p}, value at 1/3 as example, and continuity.

On the arc length sections: First a comment about voice. You have a sort of stream of consciousness literary voice. You can see this in the choice of words; for example, using "Then" rather than "Thus" creates a sense of movement through the material, rather than a sequential exposition of something. This is an attractive voice, but it is harder to control than the more neutral standard voice; it's harder to keep the reader aware of where he is in the argument and why we are doing what we are doing.

I think the Defining section should be greatly reduced, and largely replaced by a reference. This is standard stuff, adapted to defining the arc length of the graph of a function. (I am uneasy about speaking of the arc length of a function!)

The theorem should be stated early on in the next section. Your sketch of the feature of f_p that the proof makes precise is nice (though could be improved as suggested on the paper). I think you should introduce the language of good and bad at this point.

It's fine to begin with your lemma, captured in (9). But I suggest stating the lemma in that form. It's quite interesting in its own right. The y's in the statement of the lemma are exactly the x_i's. The function D (or is it M?) is totally internal to the proof of the lemma. I think you introduce it because the x_i has a hidden parameter, namely n, which you want to vary. You can do this inside the inductive proof of the lemma by calling the x_i for n-1 something different, like y_i. Or you can decorate x_i with the parameter by writing $x_{n,i}$.

The point is that if you set up a lemma, you should have it prove what you use later.

Then the probabilistic stuff starts. The reader needs some guidance. You will use Chebyshev to bound the number of "bad" values of i. Actually it's more involved than that. I would like to see a proposition stating the desired property of n which is realize by $n = \max\{N,N'\}$. The Chebyshev inequality is then invoked in the proof.

Give us guideposts along the way. What are you trying to accomplish at each step?