Re: A simple but confusing question

Source: http://sci.tech−archive.net/Archive/sci.stat.math/2004−10/0106.html

---

From: Ian Jermyn (*Ian.Jermyn_at_sophia.inria.fr*)
Date: 10/07/04

```
Date: Thu, 7 Oct 2004 14:02:49 +0200
```

There clearly is no error. The different answers just correspond to sampling with or without replacement.

One must take a bit of care in comparing these examples. The coin tossing experiment corresponds to sampling with replacement. There, as we have seen, Carroll's result does not hold, unless we do not know the value of N. However a delta function prior does of course produce the result that the probability of drawing a white is always 1/2.

In the sampling without replacement case, Carroll's result holds whether or not we know N (assuming, in the former case, that n + 1 <= N). On the other hand, a delta function prior does not produce the same result. The result with a delta function prior for sampling without replacement is that

Pr(w_{n + 1} | W_{n}, N) = (N − 2n) / (2(N − n)) ,

which is only natural: we are certain that half the balls in the bucket are white, so we cannot possibly draw more than N/2 white balls. As N becomes very large, this tends to 1/2 as one would expect, since sampling without replacement becomes equivalent to sampling with replacement.

In addition, the entropy of the distributions for the (n + 1)−th outcome, or for M, were we to calculate the probability, would be different for the two priors.

We discard the assumption p = 1/2 when the data overwhelm our prior information. How is this a paradox? In practice, we never have delta function prior knowledge, since this requires infinite precision measurements.

Ian.

```
--
--------------------------------------------------
Ian Jermyn
ianjermyn@wanadoo.fr
"George Kahrimanis" <anakreon@hol.gr> a écrit dans le message de
news:3ce8f26b.0410070306.2bb3f1ed@posting.google.com...
> Henry wrote, on 3 Oct 2004 23:38:39 +0000 (UTC):
>
```

```
> >Try to point out the error in the following
>
> There is no error there, afaIcs.
> Ian Jermyn wrote, on 5 Oct 2004 08:53:45 +0200:
>
> >I was not considering the 'sampling without replacement' case,
> >but it is simpler in fact.
>
> Here is a special case of Caroll's problem, for beginners.
> Say, we have an almost perfectly symmetric coin, so that we
> are pretty sure that it is fair in tossing. We toss it 100
> times, and it lands heads each time! What is the probability
> of the next tossing being heads? But 0.5, of course, because
> we are pretty sure that the coin is fair, so that the data does
> not matter; we rather regard the sample as a fluke.
>
> Lewis Caroll regarded the initial probability of white ball
> (during the process of stuffing the bucket)
> as fixed at 05; in the new example we have similarly fixed the
> probability of heads. If we consider other possible values of p
> besides p=0.5, we can say that Caroll's prior is a delta function
> (of possible values of p) set at 0.5.
>
> There is no mistake in Caroll's solution (and the problem can
> be answered in one line) but it is still a paradox, because
> in practice we tend to discard the assumption "p=0.5" when
> the data are stacked against it.
```