Leo's Logbook
1997

An Example of a Scientific Logbook

Scientific logbooks are bound so that the pages can not be lost or removed. The pages are numbered also, so that if the book is photocopied, it is easy to reassemble the copies in the right order. The paper is of high quality so that this record will last for many years. There have actually been cases where patent rights of considerable commercial value have been assigned in court on the basis of logbook records. Here I will track out for you a series of inquiries about a certain feature in a specific kind of plot.

Page 56, Thursday, April 17, 1997
The plot on the bottom of the page is the result of a theoretical calculation. As described in the second paragraph, I expected what we call a flat distribution, like in the plot.

Page 58, Thursday, April 17, 1997
My beautiful theory has been viciously murdered by callous and uncaring observed fact. The distribution is not flat. It is sloped, and there is a sharp peak at the left side. My first thought is that the computer program making the plot (the code) is wrong; that theory doesn't live very long. Didn't even live long enough to make it to the bottom of the page.

Page 59, Thursday, April 17, 1997
I have a new and beautiful theory about the peak at the left. I looked at it really close up and find 160 events which are at the left which have a certain property, called "zero planes hit." So I have a new hypothesis: "The peak at the left is due to zero planes hit." I know that when no planes are hit, the result is meaningless. So I get rid of the zero-planes-hit cases, and my spike goes away. At the bottom of the page, I believe I know what causes the spike, and stop thrashing away at this problem.

Page 76, Thursday, April 24, 1997
The peak at the left is back. See it there, on the second plot, labeled "data?" Darn. I thought I knew what that was, but I don't.

Page 85, Wednesday, April 30, 1997
On the bottom of the page, I make a new construction. I think that maybe the peak is from electrons. You see, this plot is made from a bunch of measurements on a bunch of particles. The assumption was "All the particles are pions"—but then I realized that if the assumption is wrong, and there are a few electrons in the bunch by mistake, then it would make a peak at the left-hand side of the plot.

Page 87, Wednesday, April 30, 1997
Here is my first test of the "Maybe some electrons snuck in" theory. If the theory is right, then the peak should be very close to zero. It should be all in that one "bin" on the very left side. Ha! It is! OK, second test . . . I think I know how many electrons could have snuck in there, dead max. Could that number of electrons make this peak?

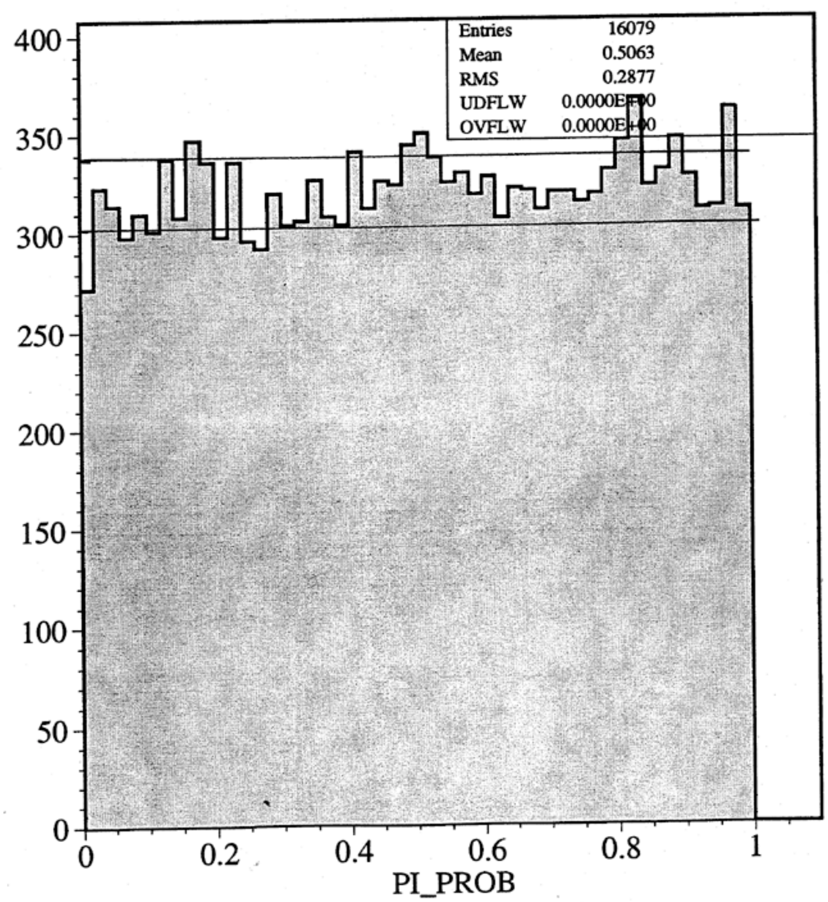
Page 88, Wednesday, April 30, 1997
Yes! I think that no more than 0.1% of the bunch of particles I call pions could actually be electrons that snuck in. And a peak of that size on the left-hand side corresponds to 0.1% contamination rate. . . . So now I think I know what causes that peak at the left.

Final note . . .
I spent two years quite certain that this peak at the left side was due to electrons passing themselves off as pions. On July 23, 1999, while working on some other thing, I saw my peak disappear . . . and I wasn't doing anything involving electrons at the time. I was playing around with some totally different thing, called (ironically enough) "accidentals." I have to wonder, "Why did I think that peak was due to electrons?" Going back over my notebooks from two years ago answers the question.

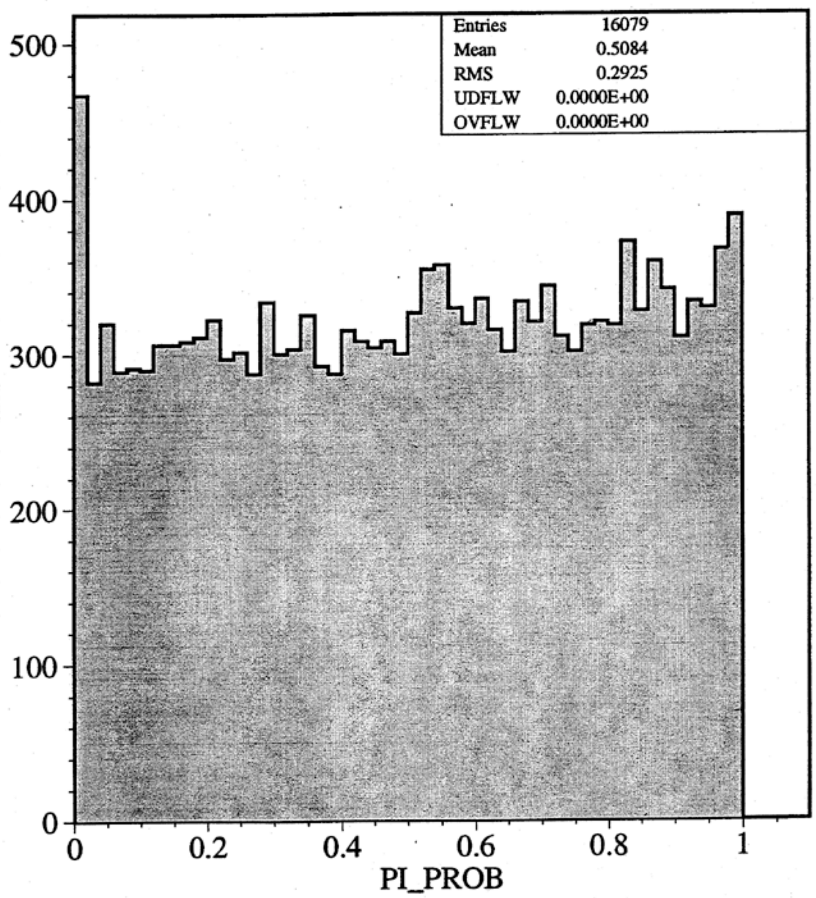
First check: does initial distribution of ADC ~~bits~~ values, when used to compute a probability distribution, give back a uniform distribution?

At first I thought it should give back a perfectly uniform distribution. The small-statistics hand-made MC calculation at right convinced me that there will be variations from perfect flatness due to the finite sample not matching a continuous distribution.

Below, the P_{COMB} distribution when a uniform random number generator, set to a specific initial seed, is used both to create and check (with the same number of events)



But using real ~~P~~ π distributions, I get instead



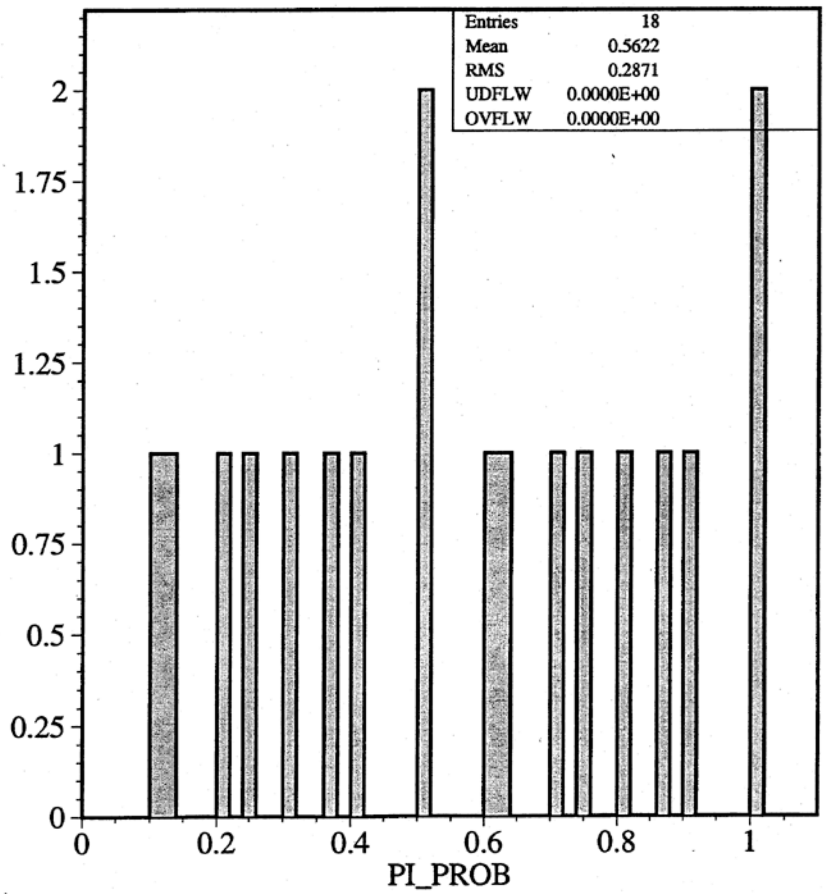
calib_selftest.hst = calib'd check on even triggers in run 8397

Because the result is flat, and for random uncorrelated numbers, I believe the code is correct and this plot shows correlations between planes in the data; the value

$$P_{COMB} = P_\pi \sum_{j=0}^{N-1} \frac{(-\ln P_\pi)^j}{j!}$$

Is proportional to the product of the probabilities, and therefore sensitive to correllations.

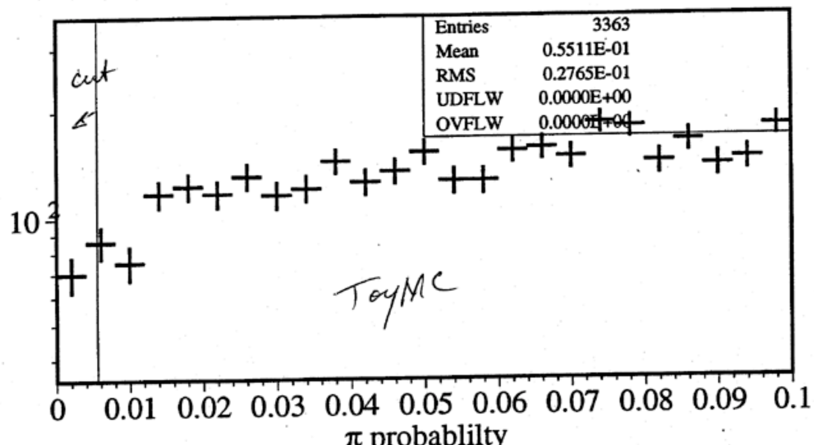
To check, I plot for the case where calib & check are done on events w/ only 1 plane hit - but there are only 18 such π in 8397:

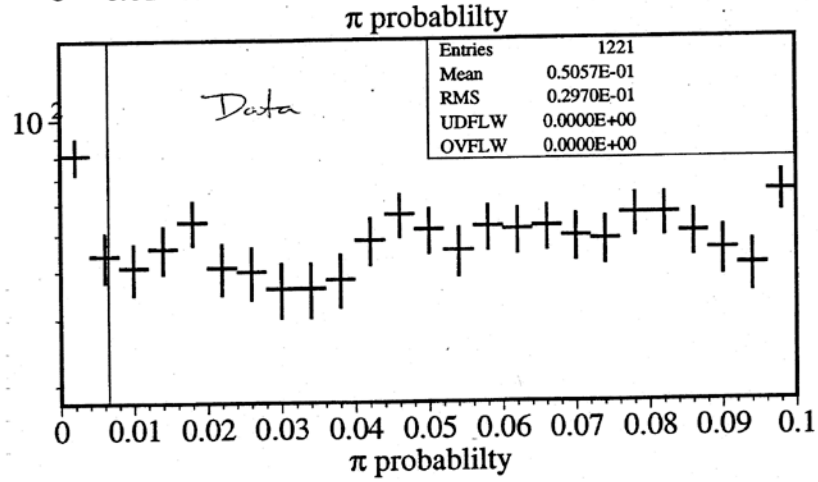


Expanding the peak near $P_{COMB}(\pi) = 0$ for the self test file, I see that ~~those~~ there are ~160 events at $P_{COMB} = 0$ exactly.

→ which all have zero planes hit.
Cut on #planes >0 & the spike goes away.
I ~~re~~make calib_self 8397.hst & stop thrashing this point.

Then I re-write the toy MC to ~~[illegible]~~ select gang2 / gang 4 once per track - but this ~~[illegible]~~ creates REJ = 875.





What are my π's that look like electrons doing?

I try removing lowest prob reading from the set of 16, but in the data this only brings REJ up to 273 or so.

30 Apr. 97

The correllations, when my random number generator is used to create ~~rando~~ ADC readings, are on the order of 2% → And greg was using a different random number generator, and he also got much better results via toy monte carlo. Wrong tree again, Fido

I modify sasha.inc, the ADC simulator, to try to get F/B correllations in overflow bins to roughly 5% level for gang 2 & 4 considered together — This is done by just setting 5% prob that back plane overflows if the front one does.

→ Well, no. counting a reading in highest calib bin for 1, all others for zero, I compute the inter-plane correllation coefficients; for the case of 16 gang: 2 readings the F/B coefficients are, in %.

2.18	1.30	2.47	1.89	$\mu = 2.87\%$
2.59	3.37	3.35	5.82	

If overflow bin is set to 1024 in back plane for (~~1.8%~~ 1.8) of the cases where the front plane is overflowing, μ comes out to be 2.88%

For the ~~back plane~~ gang 4 2 case, the data have F/B correlation coefficients, in % of

3.51	5.47	3.19	5.42	
2.23	7.49	2.11	-20.6 %	→ No correllation, even anti-correllation in ~~back plane~~ TRD8

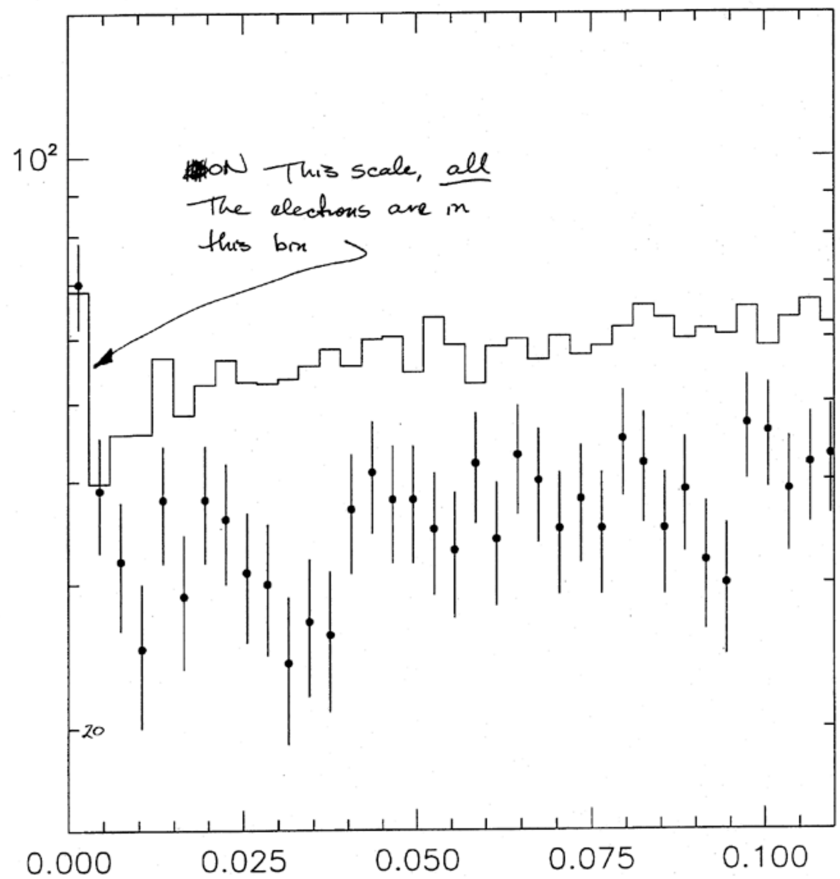
$\mu = 4.20\%$, not counting TRD8

Setting back plane to overflow in 2.80% of the cases where the front overflows for this gang-4 region gives $\mu = 4.17\%$ ... and I do no correllations for 8th TRD

I put these into SASHA.INC, re-run the calculate of constants & calibration, get rejection factor 795^{+85}_{-70} to 1.

Thinking that maybe it is real electrons in my π sample, I try Fitting ~~π~~ π_{comb} from data for 16 plane case to the toy MC curves for $e^{\pm}$, $\pi^{\pm}$ ... fit is awful & contamination looks like ~ 1 part in 500.

Same plot, expanded scale:



So there is the possibility of incorrect measurement of rejection factor due to electron contamination - what if # of passing π for that low bin is extrapolated from passing π in nearby bins?

I define extrapolated rejection factor as the one obtained by replacing the region $0 < \Pi_{COMB} < 2\times10^{-3}$ with a copy of the region $2\times10^{-3} < \Pi_{COMB} < 4\times10^{-3}$ for the π curves & using the same cut point.

	16 planes	≥ 0 planes
old	268^{+28}_{-23}	164 ± 7
extrap.	354^{+43}_{-35}	198^{+10}_{-9}

Next page.

```
Test calib on:       r08397.hstsel_10apr
Using constants in: cal2_8397.out


 Cut point is in bin   130 corresponding to prob 0.00650
  For 16 planes:
Total e=  38979   total pi=  33409    Passing pi=  112
 Rejection factor = 268. +28. -23.
Removing contamination by extrapolation, passing pi=   85
 Rejection factor = 354. +43. -35.

33409/(112-85) = e contamination in pi sample of 1 part in 1200


 Cut point is in bin   181 corresponding to prob 0.00905
  For >0 planes:
Total e=  96923   total pi=  96252    Passing pi=  529
 Rejection factor = 164. + 7.  -7.
Removing contamination by extrapolation, passing pi=  437
 Rejection factor = 198. +10.  -9.

96252/(529-437) = e contamination in pi sample of 1 part in 1000
```

Since the number of "pions" deleted by this method matches (roughly) the number of contaminating π in the e sample from the Ke3 study, I tend to believe this baby.