

Probabilities and Estimations -- a Review  
John Karl 4/3/13

Well, it sure is difficult discussing the subtle topics of probability and inference in the environment of short NavList postings, along with their kibitzing. One more try??

A. For simplicity of discussion let's limit ourselves to celestial St. Hillarie LOPs. The largest error in these is the altitude measurement. Yes there's other uncertainties, such as the nautical almanac accuracy, altitude corrections, etc., but let's include them in the altitude measurements. Then by the same observer making many altitude measurements from known locations and reasonably uniform observing conditions, a 1D probability distribution of the measurements is determined prior to setting to sea – it's prior information. Typically its standard deviation (SD) would be 1-2 nm.

B. In sailing between LOPs (or fixes) we estimate our track over the ground, or maybe just our track though the water (omitting consideration of leeway and currents). The location at the end of the first type run is normally called a ER position, at the end of the second type run, a DR position. For our discussion the distinction isn't relevant, except of course, awareness of which one is used. To avoid controversy over this terminology, I've introduced the notion of a TP, the tentative position. That's simple the ship's location at the end of a run using whatever information the navigator chose to use, but in any case, information far less accurate than celestial LOPs.

With this understanding we can now review five quite different cases:

1. The Goudsmit statement serves to point out the subtle aspect of these discussions. The statement says that if a navigator is at a fixed known location and starts plotting cocked hats (from his St. Hillarie observations), each constructed from three newly observed LOPs, he will find that each cocked hat will contain his location only 25% of the time. This is trivial to prove to yourself by sketching eight cocked hats, switching each LOP for each case to the other side of the ship's location by a symmetric amount. (The two locations for each of the three LOPs gives  $2 \times 2 \times 2 = 8$  cocked hats, and this only requires that the distributions be symmetric.)

The subtle point is the Goudsmit statement has nothing to do with navigation. It has everything to do with prior knowledge versus posterior knowledge. It will illustrate the distinction between a direct problem and an inverse problem. Goudsmit's navigator knows where he is, and his observed LOPs are centered on this known location. In navigation we're faced with the opposite case, with the inverse: We don't know where the ship is; the LOP distributions are centered on the observed LOPs, not on the ship's known location. We center these LOP probability distributions on the LOPs because we have no better knowledge – we use all we've got and no more.

So navigation is the inverse of Goudsmit – he knows where he is and computes the LOPs; navigators observe the LOPs and compute where they are. Goudsmit knows the answer and computes the question – Remember Johnnie's Carson's Great Camac? He could always give the question to the answer – that's easy. Navigators know the question and need to compute the answer – that's harder.

2. The two LOP fix is the simplest. Because we have prior knowledge of the LOP normal (Gaussian) distributions and their SD (frequently called sigma), we can calculate the probability (per area) that the ship is in any given area. Each of these distributions depends only on the perpendicular distance away from the LOP, is independent of distance parallel to the LOP, and is centered on the LOP. Since these two probs are independent, they are multiplied together to get the joint probability that the ship's

location satisfies them both. I've posted plots of this case (somewhere), showing that as expected, the MPP is at the intersection of the LOPs, and that for LOPs with unequal SDs, the major axes of the elliptical prob contour lies closer to the LOP with the smaller SD. Because the prob distributions of the two LOPs are multiplied together, the SD of the MPP is significantly greater. So the probability that the ship is less than one altitude SD from the intersection is quite small -- like a 10% chance in cases I've looked at. And that, I think, is the most valuable result of this calculation.

Finally, I must emphasize that probability predicts the relative frequency of occurrence of the same event from many trials under identical conditions. This could have some application in events like treasure hunts and ocean racing. But cruising and commercial navigators might just keep in mind that the SD for a fix is greater than for either single LOP. Counterintuitive?

3. That darned old cocked hat is now easier to appreciate from the above observations. From the product of three normal distributions, we've seen that the MPP is always inside the hat, and that it's located at the symmedian point. For practical purposes this symmedian is near the center of the hat but closest to the shortest side of the triangle and next closest to the next shortest side. We've demonstrated this in several examples in several posts.

4. The estimated position running fix (EPRF) pertains to a very different situation than cases 1, 2, and 3 above. As discussed in (B) of the introduction above, a running fix occurs when a single LOP is acquired having significantly greater accuracy than the estimation of the ship's position from less accurate means (such as logs, compasses, etc.). The significant point is that this estimation of track, yielding a TP (tentative position), has no prior determined probability distribution. The running fix is not a probability problem. It's an estimation (inference) problem. So as I said several times, the running fix is a problem in estimation, in inference: Make the best estimation possible from all the available information. Make no unwarranted assumptions and no contradictions.

Using this rationale, we acquire a LOP at a TP computed from the best information we have. We know that our TP is not as accurate as the new LOP. So rationally, we drop a perpendicular line to the LOP from the TP. This honors the accuracy of the LOP perpendicular to itself, and retains the estimation in the TP parallel to the LOP. This uses all the information in the new LOP, and retains all the useful information remaining in the TP. It uses all our info, assumes nothing, and contradicts nothing.

Now we could change this inference problem to a probability problem by inventing a TP error distribution with circular symmetry, then the EP of the running fix would still be on this perpendicular line drawn from the TP to the LOP. But the EP of the fix would be pulled toward the TP. And in addition, if we knew or wanted to estimate, the SD of the circular distribution, we could calculate where the MPP would be on this perpendicular line. (See my post #22959 for cases of both symmetric and asymmetric TP distributions.)

5. The traditional running fix (TRF) is a very different answer to this same estimation problem: The TRF assumes that the estimated DR track perpendicular to LOP1 (the last acquired LOP before the recent one) is completely accurate while, at the same time, assumes that the DR component parallel to LOP1 is completely without value. This means that the direction of LOP1 determines the direction of the estimated track that is accurate, and the direction that is worthless. How can the direction of LOP1 have anything whatsoever to do with errors in our track estimation?? Can anyone justify even one of these three assumptions??