Task Number: 2019-3

Article Title: A method for the statistical interpretation of friction skin impression evidence: Method development and validation
Date Published: 2018

Author(s):   Swofford, Koertner, Zemp, Ausdemore, Liu, Salyards

---

**Article's Subject Matter:**

- The article details the program FRStat which is being used by the US Army. FRStat is a statistical model which provides an estimate for the likelihood that two fingerprints originate from the same source.

Key Points in Article

  Explaining procedure below

- Take a latent fingerprint and an exemplar print and mark the minutiae and the minutiae angles in both prints
- Pair the minutiae in the latent and exemplar so that you minimize the minutiae distance and the difference between their angles (i.e. make them as close as possible)
- Construct the "Global Similarity Statistic" (GSS – a term created by the authors) which is simply a score representing the degree of similarity between the latent and the known. If the distances (and angles) between the two prints are smaller this will result in a higher (more positive) GSS while if the distances (and angles) are bigger this will result in a more negative GSS. GSS values around 0 are indicating an inconclusive.
- Take many cases of different source pairs for a given minutiae count between 5 and 15 and calculate the GSS for all those cases and plot their distribution (Fig 3 in paper). Take many cases of same source pairs for a given minutiae count 5 – 15 and calculate the GSS for all those cases and plot their distribution (Fig 4 in paper).
- Now for a new comparison (where the ground truth is unknown), first, calculate the GSS (we'll reference this as GSS*). Then, depending on the number of minutiae, say 9, reference the same source and different source distributions for 9 minutiae. (see picture on next page)
- Calculate from the same source distribution the probability that a GSS value would be lower than the GSS* value. From the different source distribution calculate the probability that a GSS value would be greater than the GSS* value. Then take the ratio of the same source probability and different source probability
- They conclude as follows: The latent print of exhibit # and the standard bearing the name xxxx have corresponding ridge detail. The probability of observing this amount of correspondence is approximately #### times greater when impressions are made by the same source rather than by different sources.

Actual output for prior print pair ("association")

GSS(t): 53.8689
p(GSS<=GSS(t)|SS): 0.4806
p(GSS>=GSS(t)|DS): 0.00001467
Ratio: 32761.9841
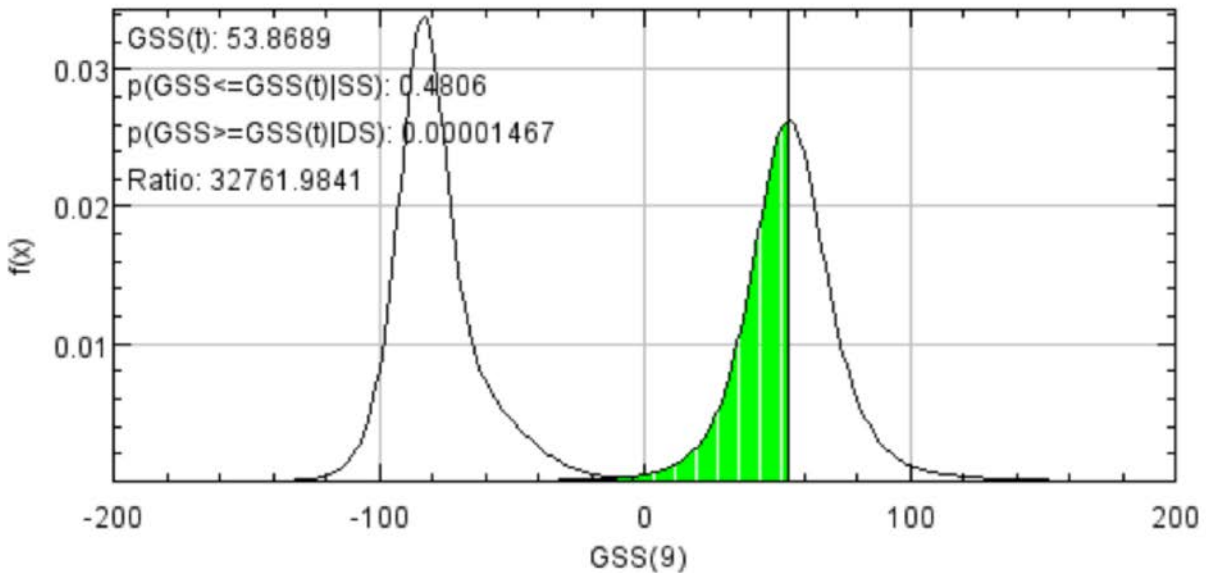
Fallacies and or Issues

- My main concern with all statistical models is the validation. I have no doubt that they perform very well generally speaking, however, when providing a risk estimate (probability, likelihood, etc) you need to make sure those "probabilities" are reliable. That is the theoretical values provided by the model should converge to the frequencies they are representing in a large scale study. I have not seen a validation study for these models that I am comfortable with.
- I am growing more and more concerned about people's (especially the lay person) ability to understand statistics. Even very simple statistical methodology seems to get misconstrued and misused. These models are even more complicated, so integrating them into our current court system would be challenging.