NOTES 7: On-Source Analysis, R.O.C. Curves, and G.R.B's

R.O.C. - Curve : detection prob., vs. false alarm prob   stands for R.eceiver O.peration C.haracteristics

On-Source analysis, Proper: We create a single noise file with multiple injections.

On-Source analysis, Our Version:

Begin by building a set of data waveforms with noise

1. Specify the # of data sets you wish to test (N, index i), then create the N random wave forms as follows:
   for A*, sample [1, 50]   for f*, sample [90, 110]   for γ*, sample [0, 2]   for $t_0 = t*$, sample [0, $t_f - (Len)$]

   We constructs signal a $h(A*, f*, \gamma*)_i$, and then $d_i(t) = n(t)_i + h(A*, f*, \gamma*)_i$

Separately, create a set of wave-templates over parameter-space (A, f, γ) / (f, γ) ($N_A \cdot N_f \cdot N_\gamma$, index j).

2. Loop through this space and create base statistics $m_{i,j}(t)$, and $\chi^2_{i,j}(t)$;
   use $d_i(t)$ as input 1, and the current $h^T_j(t)$ as input 2 with your $d_i(t)$ and some separate $h^T_j(t)$

3. Store $m_{i,j}(t)$, $\chi^2_{i,j}(t)$ and $d_i(t)$ as attributes

Now we choose some statistic and examine its efficiency; its a function of $m_{i,j}(t)$, $\chi^2_{i,j}(t)$; label it $\rho_{i,j}(m,\chi) = p_{i,j}(t)$.

5. Generate the 'foreground' values: $\max_{[t_0-\Delta, t_0+\Delta]} (\rho_{i,j}(t))$, where t_0 is the point of injection of the waveform in $d_i(t)$

6. Generate the 'background' values: step 2, but maximize over each time intervals that's not the 'foreground values' window

7. save the values into an array of the following form:

output = [[[$p_{i,1}(t)$ off source] , ... , [$p_{i,\, N_A \cdot N_f \cdot N_\gamma}(t)$ off source]]   ,   [$\rho_{i,1}(t)$ on source, ..., $\rho_{i,\, N_A \cdot N_f \cdot N_\gamma}(t)$ on source]   ]

where '$\rho_{i,j}(t)$ on source' are scalars   and   '[$\rho_{i,j}(t)$ off source]' are arrays

The above can be done within the constructor (or better?), with multiple generated-datas; the following is separate:

7. 'false alarm probability': Now for a given $p_{i,j}(t)$, we calculate $R_{i,j} = N_{BG}(\rho > \rho_{on-source})/N_{BG}$

   N_{BG} = # of background subdivision $\approx \lfloor T/2\Delta \rfloor$

   $N_{BG}(\rho > \rho_{on-source})$= number of background values greater than the foreground value

8. now, choose some threshold value T; you will compare your on-source interval rho and all your $N_{BG}(\rho > \rho_{on-source})$ rhos and see if any go over rho.

Now, the above is all for a single set of generated data checked against multiple possible templates. What we want now is to check if this single test was successful. To do so, use the chosen threshold value T, and look at the set of on/off-source's which go over this threshold for each individual template-check. If *ANY* of the template-checks on this data have an on-source value which corresponds to the largest rho value over T (within the template-checks set of on/off sources), we count the data-test as successful.

After doing this test for one set of generated-data, we repeat it for multiple sets of generated data; we sum the number of generated-data's which had successful tests at a given threshold, and label this sum $\#N(T)$. we then plot T vs. $\#N(T)$.

Lastly, we can create a heat map over an $f\ vs.\gamma$ plane by counting the 'height' at a given (f, γ) pair as the number of successful generated-data tests which has (f, γ) as their randomly generated-data f and gamma