variables.mat:

- InputSpectrumPtR: smoothed average Pt. Reyes spectrum [omega, M]
- MphiK1-MphiK6: transfer function of L_1-L_6 [magnitude, phase in radians]
- PtReyesBuoyData
    - PtReyesBuoyData.time: time stamp YYYYMMDDHHMMSS
    - PtReyesBuoyData.z: wave elevation in cm, each row a record
- rmsDLG: rms of derived processes L_1-L_6, from DLG
- rmsPtReyes: rms of derived processes L_1-L_6, from Pt. Reyes buoy data
- pre-run files for looking at wave group statistics:
    - for considering DLG waves based on Pt. Reyes TEV, 30-min:
        - cdfInfo_ptrTEV30min
        - DLGstatsHptrTEV30min
        - DLGptrTEV30min
    - for considering DLG waves based on theory TEV, 30-min:
        - cdfInfo_theoryTEV30min
        - DLGstatsHtheoryTEV30min
        - DLGtheoryTEV30min
    - for considering DLG waves based on theory TEV, 25-hr:
        - cdfInfo_theoryTEV25hr
        - DLGstatsHtheoryTEV25hr
        - DLGtheoryTEV25hr
    - MCS: 30-min (same TEV as DLG waves based on theory, TEV 30-min)
        - MCS30min.kj: $\left[\frac{\widehat{L_j}}{\sigma_j}, index\ location\ of\ \widehat{L_j}\right]$
    - MCS: 25-hr (same TEV as DLG waves based on theory, TEV 25-hr)
        - MCS25hr.kj $\left[\frac{\widehat{L_j}}{\sigma_j},\ index\ location\ of\ \widehat{L_j}\right]$
- Can directly run:

```
assembleNLDLGjointPDF(option,ka,kb,DLGptrTEV30min,DLGtheoryTEV25h
r,zkHatData,MCS25hr,DLGtheoryTEV30min,MCS30min)
```

    - option 1: compare DLG waves based on Pt. Reyes TEV & Pt. Reyes buoy data
    - option 2: compare DLG waves based on theoretical 30-min TEV & MCS
    - option 3: compare DLG waves based on theoretical 25-hr TEV & MCS
    - ka: wave group index of interest
    - kb: another wave group index of interest

Files which are indirectly used in probabilistic framework:

- Bell.m (calculate the Bell number)
- SetPartition.m (calculate set partitions- the possible maxima clusters)
- Stirling2nd.m (calculate Stirling number of the 2nd kind)
- generateConsistentCharNumber.m: to generate character string to load in DLG files

To run files on your own:

1. `zkHatData = calcZkHat(PtReyesBuoyData,rmsPtReyes)`
    - this goes through the Pt. Reyes buoy data and gets the info on the wave groups and filters out bad time series
    - `zkHatData.kj` gives info on wave groups of j waves, each row a time seroes
    - $\left[\frac{\widehat{L_j}}{\sigma_j}, index\ location\ of\ \widehat{L_j}, H_s, T_p\right]$
    - `zkHatData.zj` gives time series of the derived process, L_j in [m]
    - note that zkHatData.z1 is the wave elevation record itself in [m]

2. choose which folders (30-min TEV based on Pt. Reyes '`SeaState/30min/z#P'`,30-min TEV based on theory '`SeaState/30min/z#'`, or 25-hr TEV based on theory '`SeaState/25hr/z1P')`
   - right now, this runs through the DLG time series saved in these folders. If you use a different RCWT, set up files the same way
3. `cdfInfo = getCDFinfo(MphiK1,MphiK2,MphiK3,MphiK4,MphiK5,MphiK6,rmsDLG)`
4. `DLGstatsH = DLGstatsHOLD(MphiK1,MphiK2,MphiK3,MphiK4,MphiK5,MphiK6,rmsDLG,cdfInfo)`
5. `DLGstats = DLGstatsPLACE(DLGstatsH,cdfInfo)`
6. `[ffInfo,masterList,numT,chosenInfo,probCount] = getNLDLGts(DLGstats)`
7. `assembleNLDLGjointPDF(option,ka,kb,DLGptrTEV30,DLGmcsTEV25hr,zkHatData,PtRmcs25tev,DLGmcsTEV30,PtRmcs30min)`
   - option 1: compare DLG waves based on Pt. Reyes TEV & Pt. Reyes buoy data
   - option 2: compare DLG waves based on theoretical 30-min TEV & MCS
   - option 3: compare DLG waves based on theoretical 25-hr TEV & MCS
   - ka: wave group index of interest
   - kb: another wave group index of interest
8. `[ffInfo,masterList,numT,chosenInfo,probCount] = getNLDLGts(DLGstats)`
   - ffInfo is a struct which gives the different possible maxima configurations
     - each field is a possible maxima config. (e.g. z12345 means surrogate processes 1-5 cluster together; z1z2345 means surrogate process 1 occurs separately over an exposure while surrogates 2-5 occur clustered etc.)
     - within each field, we have the probability of experiencing that maxima cluster (pz...), the info on the time series which contain those clusters (z... - organized the same way as DLGstats: marker zA ZB ...), and the chosen indices based on numT, the probability of experiencing that maxima config., and the distribution of those maxima clusters
   - masterList gives the indices of all chosen DLG time series based on numT (note length(masterList)>=numT because some time series contain un-clustered maxima, and each index of MasterList is a time series which may excite clustered or un-clustered surrogate maxima
   - numT is the max number of NL-DLG time series which can be assembled without oversampling the maxima cluster distributions
   - chosenInfo is a struct which gives similar info as ffInfo but removes maxima configurations where the probability of experiencing that maxima configuration is so low that we wouldn't sample any of these time series