Boot strap example discussion for the lightning data case is given on p.p. 168 – 169.

What the <u>less</u> res<u>trictive</u> null hypothesis can account for in bootstrapping.

- (spread does not necessarily just depend on seeding
- other aspects of the distributions may be different.

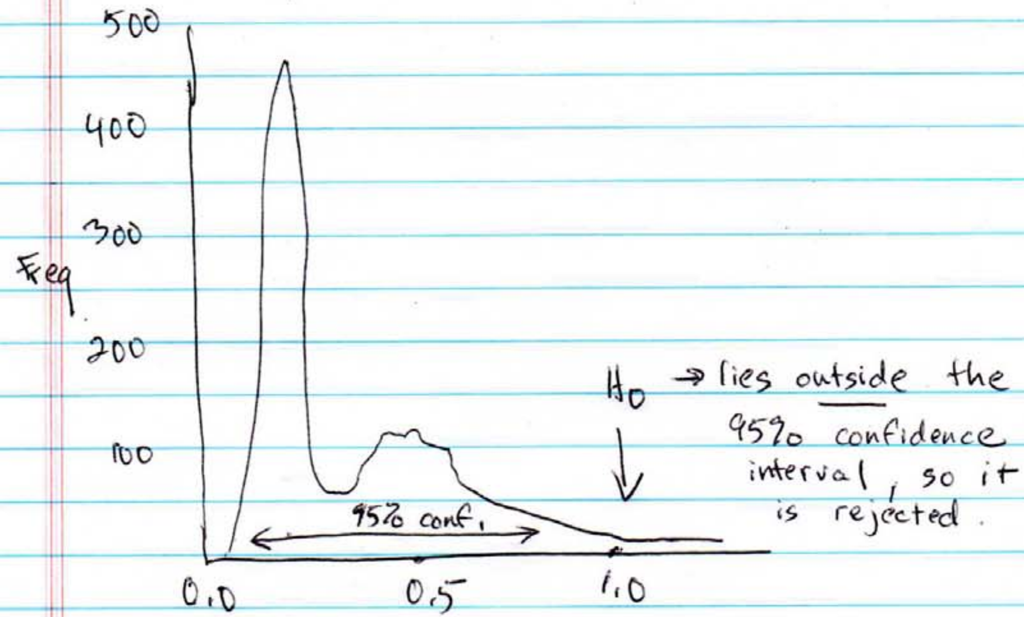


Fig. 5.9.

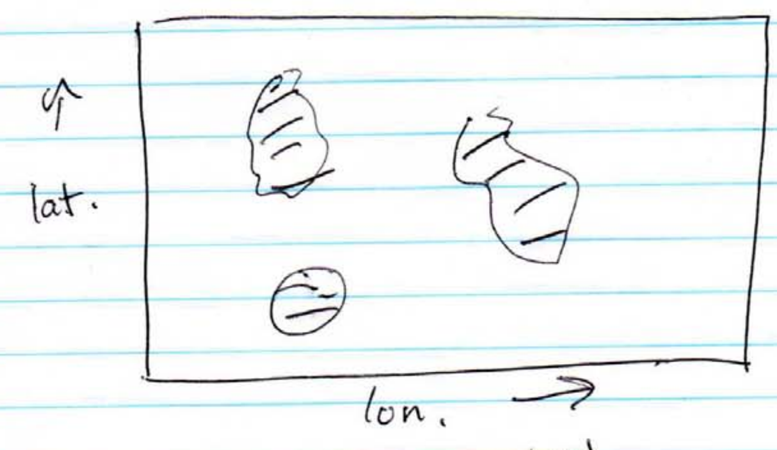
∴ other aspects of the distribution are probably different!

Field significance - A VERY important non-parametric test in geophysical data analysis.

"The" reference for this, which has been cited thousands of times, is Livezey and Chen (1983). (And if you don't do it, Livezey may embarass you at a conference!!)

In geophysical fields, often looking at patterns of significance in space, based on many local tests at the individual gridpoints.

Say you're doing a composite analysis. and you get a map which indicates the level of significance at each point.



Shaded areas represent local significance at the gridpoints exceeding some critical value. Say the 95% level on a t-test.

Two things you need to think about before you conclude things are statistically significant or not.

1) Are the data related in time? That is, is there autocorrelation present? If so, this will reduce the degrees of freedom used in the local test.

We'll address how to handle this a bit later in the course --

For now, we'll assume that the data are independent in time.

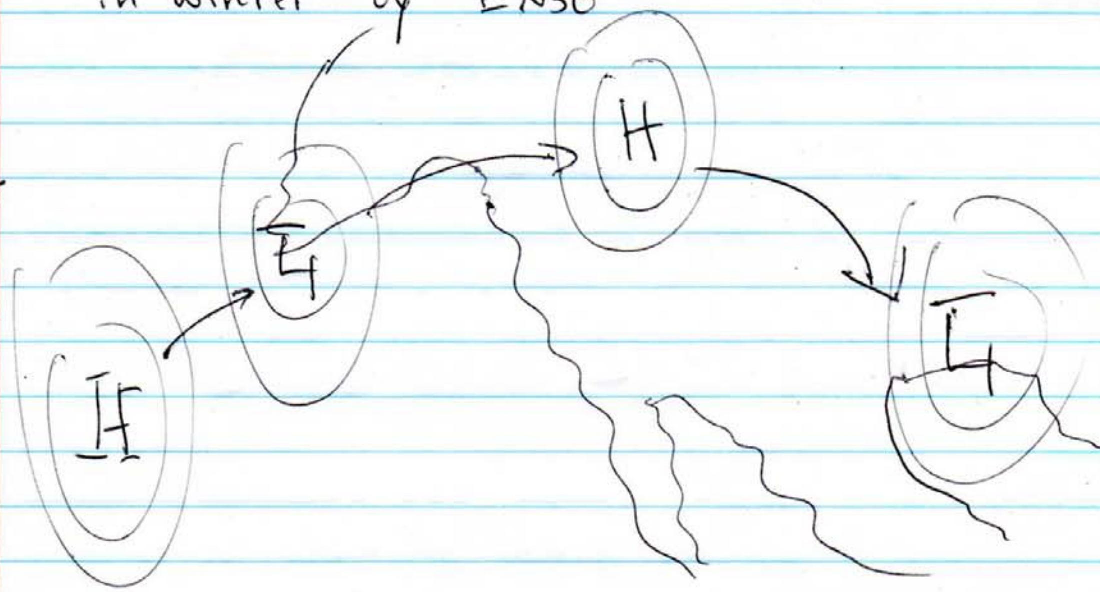
2) Are the data related in space? If so, then it would be expected that just a few dominant spatial patterns ~~explain~~ account for the significant shaded areas on the map.

Example of spatial patterns: 700-mb heights over the northern hemisphere.

The geopotential height at one point is more often than not related to what happens at other points.

Consider the PNA pattern, which get if composite 700-mb heights in winter by ENSO

PNA

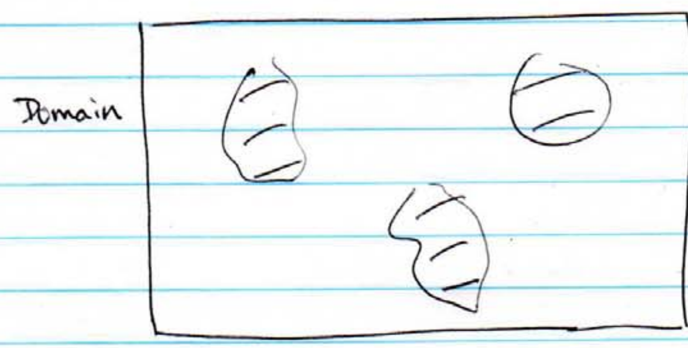


Result of changes in convection in the eastern tropical Pacific → produces a hemispheric pattern of height anomalies, or teleconnection.

So because there are teleconnections, or relationships in space, what happens at one point IS NOT independent of what happens at another point – there is spatial correlation!

How to account for this?

Original composite based on n, from a sample of n (e.g. ENSO years)



Procedure to assess significance of shaded regions in the composite example.

1) Calculate the percentage of area within the domain that passes the local significance test.

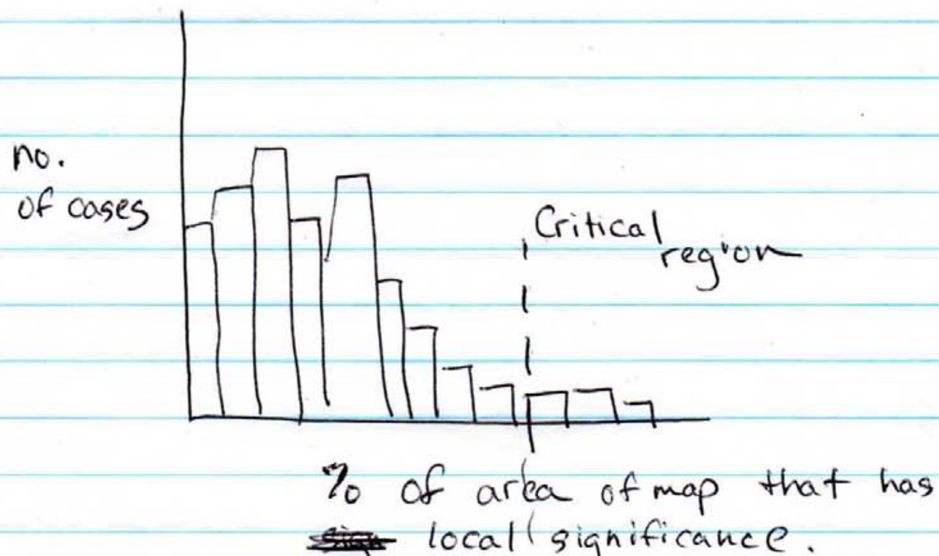
2) ~~Re~~ Randomly reorder the maps

3) Calculate the percentage of area

that passes the local significance test.

4) Repeat this procedure many times to generate a null distribution of percentage area on the map that satifies a local test.

Fig. 5.12 from Wilks.



5) **Pass test**

<u>IF</u> the percentage area on your original map falls within the critical region, <u>then field significance</u> is satisfied.

Fail test

<u>IF</u> the percentage area on original map is below the critical region, then no field significance. In other words, can get the same pattern of sig. shaded regions by a random reordering of the maps.

In terms of how to program

Have a three dimensional array for a given variable:

In FORTRAN

$VAR(i, j, t)$

x-dim ↗ (i), y-dim ↑ (j), time dimension ↖ (t)

To evaluate field significance:

- Keep x-dim & y-dim the same
- Only reshuffle the maps by changing the ~~t~~ time index.