Brief note on degrees of freedom.

Degrees of freedom = # indep. samples − numbers of parameters in stat. to be est.

t - stat. ($\nu = N-1$) $\qquad \dfrac{\bar{X} - \mu}{\frac{s}{\sqrt{N-1}}}$ $\qquad \mu$ est. (1 param)

χ^2 ($\nu = N-1$) $\qquad \chi^2 = \dfrac{(N-1)\, s^2}{\sigma^2}$ $\qquad \sigma^2$ estimated (1 param)

CAVEAT IN DETERMINING DEGREES OF FREEDOM!!

- The samples N <u>must</u> be "independent"
- Independence means samples are not related, or autocorrelated, in <u>time</u> and <u>space</u>.

More on this later ...

<u>Hypothesis testing</u>

What are we trying to do?

Evaluate whether something satisfyes a null hypothesis (H_0) or an alternative (H_1)

H_0 = ~~is~~ statistically insignificant
H_1 = statistically significant (i.e interesting)

In statistics, the word "significant" means that the alternative hypothesis is satisfied :- a very specific definition!!

Typical significance levels = 90%, 95%, 99%

Type I error: Reject hypothesis and it turns out to be correct

Type II error: Accept hypothesis and it's wrong.

Lect. 5

How to do hypothesis testing:

1) State the desired significance level
2) Formulate the null hypothesis and the alternative hypothesis.
3) Choose the statistic (e.g. t-stat)
4) Define critical region
5) Evaluate the statistic and state the conclusion
   "H_0 is rejected with x% certainty.

Hartmann's simple example: Dilbert's ski vacation:

- Dilbert skis 10 times per season
- Temperature avg. on days there is 35°F
- Climatological mean is 32°F
- standard deviation = 5°F

1) use a significance level of 95%
2) <u>Null hypothesis</u>: The mean for Dilbert's 10 days is no different than climatology
   <u>Alternative</u>: Days were warmer than climatology
3) use t-statistic

two-sided $t = 2.26$

$$t = \frac{\bar{x} - \mu}{\sigma}\sqrt{N-1} = \frac{35-32}{5}\sqrt{9}$$

4) $t = 1.80$ with 9 deg. of freedom.
   $t_{critical} = t_{0.05} = 1.83$ (one-sided)

∴ t-test is not satified.

5) Conclusion : Dilbert did not experience conditions that were statistically ~~significantly~~ significantly different from the mean climate. (ie. null hypothesis is not rejected).

Compositing or superimposed epoch analysis.

Idea: Isolate signal of an event from the background by averaging data in relation to the event.

Advantage: ~~[illegible]~~ Simple and powerful technique. Statistics are very straightforward to interpret.

Disadvantage: Can be easily misused. ~~How the~~ What criteria are used to select composite ~~is selected~~ is at the discretion of the user.

Steps in compositing

1) Determine categories
2) Compute mean, variance, etc. for each category
3) Statistical test
4) Display results with significance.

Significance of Composite results

1) Only use the ~~quantitative~~ statistical tests if you expect the result "a priori" — or in other words there is not a "large" relationship in time and space in the data (i.e. autocorrelation)

We'll talk about ways to account for this later after we talk about correlation.

2) You should subdivide the sample to make sure data are independent (more later..)

3) Composites should be based on a good theory that has an objective physical basis (e.g. ENSO, diurnal cycle, etc..) That is a good "a priori"

If there is no a priori expectation in constructing the composite, the relationship was found after "fishing."

Some examples of Compositing — and their misuse (Hartmann's notes)

ENSO Compositing example.

1) Composite all the months in a 50 year record according to Niño 3 index. ≈ 600 months.

   100 months in a positive composite (El Niño)
   100 months in a negative comp. (La Niña)

2) Compute mean and standard deviation of monthly Tucson rainfall of each composite, then use a

difference of means test to evaluate whether the El Niño and La Niña composites are statistically different or not.

~~3) Get the result that~~

Question: Should we trust the results from this statistical analysis? Why or why not?

This example makes two errors

1) There is autocorrelation in the ENSO data. That is, SSTAs in the tropical Pacific are not going to change that much over a month. So the true number of independent samples should take this into account is actually much less. → "A posteriori problem"

2) The formulation of the hypothesis is not based on a very good physical basis – do we assume ENSO relationships and physical mechanisms of rainfall are constant all years in AZ?
→ "A priori problem"