Example like what will be used in homework assignment:

- Have approximately 50 years of winter precipitation data in Tucson (Dec - Mar)
- Use Niño 3.4 index to classify ENSO pos. or neg. winters using some threshhold for the index
- Say get 12 ENSO+ winters, 13 ENSO− winters.
- Perform a difference of means test to see if the average precip. for each composite is different.
- Any idea what the answer might be?? Why?

We can also test whether variances are different.

Chi - Squared Dist. - Test of Variance. (χ^2)

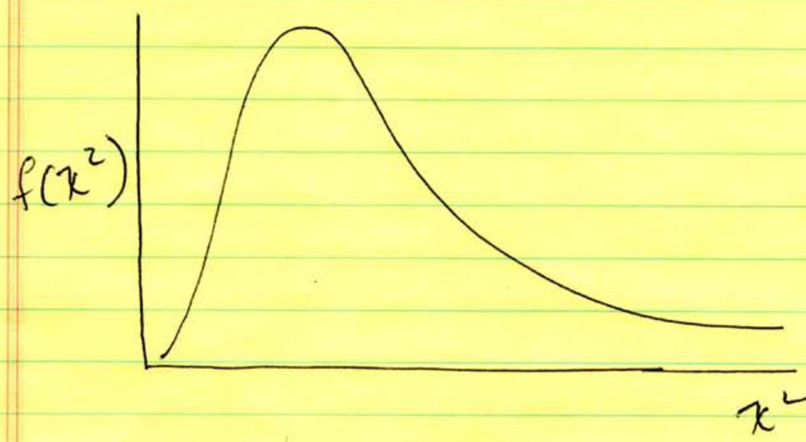
This is a special case of the gamma distribution (we'll talk about that one later)

$$f(\chi^2) = f_0(\nu)\,(\chi^2)^{(\frac{1}{2}\nu - 1)}\, e^{-\frac{1}{2}\chi^2}$$

$$\chi^2 = (n-1)\frac{s^2}{\sigma^2}$$

ν = degrees of freedom.

This distribution is positively skewed:



Like t-distribution, χ^2 is function of the degrees of freedom:

Test for confidence interval (e.g. 95%)

$$\frac{(N-1)s^2}{\chi^2_{0.025}} < \sigma^2 < \frac{(N-1)s^2}{\chi^2_{0.975}}$$

Read these values off a table.

To compare variances of different samples from the same population, ~~can~~ use the F-statistic.

$$F = \frac{s_1^2}{s_2^2}$$ with N_1 and N_2 ~~degrees~~ being the sample sizes.

$$\left.\begin{aligned} v_1 &= N_1 - 1 \\ v_2 &= N_2 - 1 \end{aligned}\right\} \text{degrees of freedom.}$$

Relationship to χ^2 test: F statistic is the same as χ^2 with $v_1 = N-1$ and $v_2 = \infty$.

These distributions are important when doing spectral analysis, as we'll see later in the course.

Brief note on degrees of freedom.

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

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

χ^2 ($v = N-1$) $\qquad \chi^2 = \dfrac{(N-1)\, s^2}{\sigma^2} \qquad$ σ^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 satisfies 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) Null hypothesis: The mean for Dilbert's 10 days is no different than climatology
   Alternative: 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 <u>not</u> satified.

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

<u>Compositing</u> or superimposed epoch analysis.

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

<u>Advantage</u>: ~~Simply~~ Simple and powerful technique. Statistics are very straightforward to interpret.

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

<u>Steps</u> 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

~~[illegible]~~

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 <u>two</u> <u>errors</u>

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 year in AZ?
→ "A priori problem"

Another example of a priori error.

- Rainy days in Seattle in December.
- Calculate mean and std. dev. for each day
- Compare this to the monthly mean.
- If you find one day that is statistically sig. different, you must consider the probability of <u>every</u> day being significant - even though the data are independent.

Probability that none of the 31 days will pass 99% significance

$$(0.99)^{31} = 73\%.$$

Can also use composite analysis to compare variances in the F-test:

<u>Example</u>: Does QBO impact variance of T in the NH polar vortex.

40 yrs. of Jan. mean data:

16 winters $\Rightarrow$ east

16 winters $\Rightarrow$ west. $\quad n_1 = n_2 = 16$

$s_1 = 12K$

$s_2 = 7.5K$

$$F = \frac{s_1^2}{s_2^2} = \frac{(12K)^2}{(7.5K)^2} = 2.56$$

Critical value for F-test = 3.52

$\therefore$ Conclude it is not significant.

df $\xrightarrow{N_2}$

F-table.

$N_1 \downarrow$

Sig. values.