TIME SERIES (Chapter 8 of Wilks)

In meteorology, the order of a time series matters!
We will assume stationarity of the statistics of the time series. If there is non-stationarity (e.g., there is a diurnal cycle, or an annual cycle) we will subtract the climatological mean and standardize the anomalies: $z(t) = \frac{x(t) - \mu(t)}{\sigma(t)}$. Otherwise, we can stratify the data (e.g., consider all winters together as a single time series).
Time series can be analyzed in time or in frequency domains.

Time series models

Example of a simple time series model
$x_{t+1} - \mu = \phi_1(x_t - \mu) + \varepsilon_{t+1}$. This is an autoregressive model of order 1 (AR(1)). Such model can be used to:

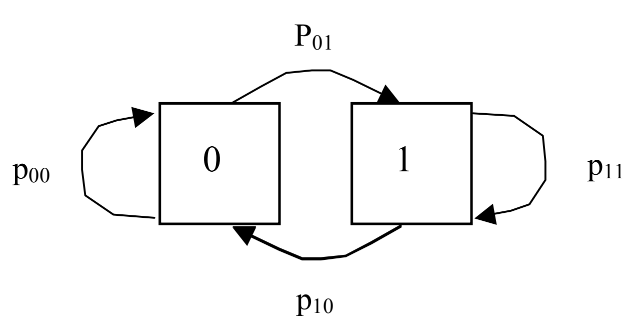
a) Fit the time series and derive some of its properties. Similar to fitting a theoretical probability distribution to a sample.
b) To make a forecast: $\hat{x}_{t+1} - \mu = \phi_1(x_t - \mu)$

For discrete time series, the equivalent of autoregressive models are Markov chains.

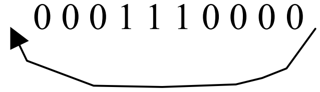
Discrete time series (Markov chains)

We have a discrete number of states (e.g., 2 or 3 states). For example, rain:1, no rain: 0. At a given time, the chain can remain in the same state or change state (MECE).

Example: 2-state first order Markov chain



For example: a periodic time series 0 0 0 1 1 1 0 0 0 0



Transition probabilities: $p_{01} = P(x_{t+1} = 1 \mid x_t = 0)$ etc.

In the example above: $p_{01} = \frac{1}{7}; \quad p_{00} = \frac{6}{7}; \quad p_{10} = \frac{1}{3}; \quad p_{11} = \frac{2}{3}$

Note that $\begin{array}{l} p_{01} + p_{00} = 1; \\ p_{10} + p_{11} = 1 \end{array}$ both are MECEs.

We can also define unconditional probabilities: $\begin{array}{l} \pi_0 = P(x_{t+1} = 0) \\ \pi_1 = P(x_{t+1} = 1) \end{array}$

$$\pi_0 = \frac{7}{10} = \frac{p_{10}}{p_{01} + p_{10}} = \frac{1/3}{1/7 + 1/3}$$
$$\pi_1 = \frac{3}{10} = \frac{p_{01}}{p_{01} + p_{10}} = \frac{1/7}{1/7 + 1/3}$$

These represent

$$\text{probability of being in state 0 (or 1)} = \frac{\text{probability to change to 0 (or 1)}}{\text{probability to change}}$$

(See demonstration later)
We can also define

Persistence = lag-1 autocorrelation = $r_1 = corr(x_{t+1}, x_t)$

$r_1 = p_{11} - p_{01} =$ probability of being in 1 coming from 1, minus probability of being in 1, coming from 0.

$$r_1 = \frac{2}{3} - \frac{1}{7} = \frac{14-3}{21} = \frac{11}{21}$$

We can also compute it as $r_1 = p_{00} - p_{10}$:

$$r_1 = \frac{6}{7} - \frac{1}{3} = \frac{18-7}{21} = \frac{11}{21}$$

Note: persistence implies that $p_{01} + p_{10} \le 1, \quad p_{00} + p_{11} \ge 1$, i.e., there is a stronger tendency to remain in a state than to change states ($p_{00} + p_{11} > p_{01} + p_{10}$).

Then from $\pi_1 = \frac{p_{01}}{p_{10} + p_{01}}$, we get that $p_{01} \le \pi_1$, and similarly $p_{10} \le \pi_0$ (i.e., if there is persistence, the probability of transitioning into a state from the other is smaller than the unconditional probability of being in that state).

Furthermore, $p_{10} = 1 - p_{11} \le \pi_0 = 1 - \pi_1$, or $\pi_1 \le p_{11}$.

In summary, if there is persistence, $p_{01} \le \pi_1 \le p_{11}$ and $p_{10} \le \pi_0 \le p_{00}$.

Exercise: show that $\pi_1 = \frac{p_{01}}{p_{01} + p_{10}} = \frac{n_1}{n_1 + n_0}$

Proof of this:

$\pi_1 = \frac{p_{01}}{p_{01} + p_{10}} = \frac{n_{01} / n_0}{n_{01} / n_0 + n_{10} / n_1} = \frac{n_{01} n_1}{n_{01} n_1 + n_{10} n_0} = \frac{n_1}{n_1 + n_0}$ since $n_{01} = n_{10}$ because on the long run, there have to be as many changes from 0 to 1 as the other way around.

Actually $\pi_1 = \frac{n_1}{n_1 + n_0}$ is a more natural definition of unconditional probability, so the proof really shows that $\pi_1 = \frac{p_{01}}{p_{01} + p_{10}}$ as defined before.

Exercise: compute r_1 from the lag-1 autocorrelation.

Hypothesis testing for the presence of persistence in the time series:

	$x_{t+1}=0$	$x_{t+1}=1$	
$x_t=0$	6	1	7
$x_t=1$	1	2	3
	7	3	10

Marginal totals

Actual numbers from the series 0 0 0 1 1 1 0 0 0 0

Now the null hypothesis is that there is no persistence, so we create the table corresponding to no persistence:

	$x_{t+1}=0$	$x_{t+1}=1$	
$x_t=0$	4.9	2.1	7
$x_t=1$	2.1	0.9	3
	7	3	10

Marginal totals

For example

$$4.9 = \pi_0 \pi_0 n = \frac{7}{10} * \frac{7}{10} * 10$$

So, to test whether a series really has $r_1 \neq 0$ (and it's not just sampling), we can use the X^2 distribution with 1 degree of freedom, since the marginal totals are given from the sample, so that given a single value on the table, the others are determined.

Null hypothesis $r_1 = 0$. The hypothesis of independence (columns are independent of the rows) is tested with

$$\mathrm{X}^2 = \sum_{classes} \frac{(\#\text{observed} - \#\text{expected})^2}{\#\text{expected}}$$ (see note below), so that

$$\mathrm{X}^2 = \frac{(6-4.9)^2}{4.9} + \frac{(1-2.1)^2}{2.1} + \frac{(1-2.1)^2}{2.1} + \frac{(2-0.9)^2}{0.9} = 2.74$$

Since the 5% X^2 for 1 d.o.f is 3.84, the persistence of the time series we created is not significant at a 95% level: we could have obtained a size 10 sample with such apparent persistence even though the population has no persistence with a probability greater than 5%.

<u>Note: Reminder from class 3:</u>

An important use of the chi-square is to test **goodness of fit**:

If you have a histogram with n bins, and a number of observations O_i and expected number of observations E_i (e.g., from a parametric distribution) in each bin, then the goodness of fit of the pdf to the data can be estimated using a chi-square test:

$$X^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$ with n-1 degrees of freedom

The null hypothesis (that it is a good fit) is rejected at a 5% level of significance if $X^2 > \chi^2_{(0.05,n-1)}$. The table above has 4 bins but only one d.o.f.

The table above is known as **contingency table.**

	$x_{t+1} = 0$	$x_{t+1} = 1$	
$x_t = 0$	6	1	7
$x_t = 1$	1	2	3
	7	3	10

Marginal totals

We were checking whether the Markov chain has persistence, i.e., whether the value at *t+1* is dependent on the value at *t*.

Example of a test of independence in a contingency table.

	democrat	republican	independent	
Women	68	56	32	156
Men	52	72	20	144
	120	128	52	300

Marginal totals

A test of independence checks whether the null hypothesis that political affiliation is independent of gender is valid. The null hypothesis would generate a table like

	democrat	republican	independent	
Women	62.40	66.56	27.04	156
Men	57.6	61.44	24.96	144
	120	128	52	300

Marginal totals

where, for example, the number of democratic women is obtained as “prob. of being a woman* prob. of being a democrat (gender independent)* number of people surveyed” = (156*300)* (120/300)* 300=62.4.

Since the marginal totals are fixed, the number of dof for each row is (c-1)=(3-1)=2, and the number of dof for each column is (r-1)=(2-1)=1, so that the total number of dof in this contingency table is (c-1)(r-1)=2*1=2. Here r is the number of rows and c the number of columns.

Consider the first column, democrats. If we accept the null hypothesis that the value of p=62.4/120 is the gender independent probability that of this group of people, democrats are equally distributed among women and men, this is a binomial distribution with an expected value $np = 62.40$. (The expected value for men is $n(1-p) = 57.6$).

Suggestion of a demonstration of the test of independence:
Only for one column, a binomial distribution, e.g., the probability of being a woman or a man if you are a democrat:

Consider the test statistic for this binomial distribution:

$$T = \frac{(68-62.4)^2}{62.4} + \frac{(52-57.6)^2}{57.6} = \frac{(X_1-np)^2}{np} + \frac{(X_2-n(1-p))^2}{n(1-p)} = \frac{(X_1-np)^2}{np(1-p)}$$

where we have used $X_2 = n - X_1$ and $\frac{1}{p} + \frac{1}{1-p} = \frac{1}{p(1-p)}$

The variance of the binomial distribution is $np(1-p)$.

Therefore,

$$T = \frac{(X_1-np_1)^2}{np_1} + \frac{(X_2-np_2)^2}{np_2} = \frac{(X_1-np_1)^2}{np_1(1-p_1)}$$

In other words, T is the mean square of a random (binomial) anomaly divided by its variance, and for large *n*, when it approaches a normal distribution, this ratio has approximately a X^2 distribution with one degree of freedom.

In this case T=1.05, and for $X^2_{0.05,1}$=3.841 so just knowing that of a group of 120 democrats 68 were women does not show that women tend to vote democrat.

For several columns, this generalizes to

$$T_{(r-1)(c-1)} = \sum_{\substack{i=1,r \\ j=1,c}} \frac{\left(O_{ij} - E_{ij}\right)^2}{E_{ij}} \sim X^2_{(r-1)(c-1)}$$, which is the test that we have used above.

In this case, T=6.43, whereas $X^2_{.05,2} = 5.99$. This shows that the null hypothesis of independence between rows and columns can be rejected with 95% confidence, and political affiliation is gender dependent.

<u>End of note</u>

Uses of the Markov chain:

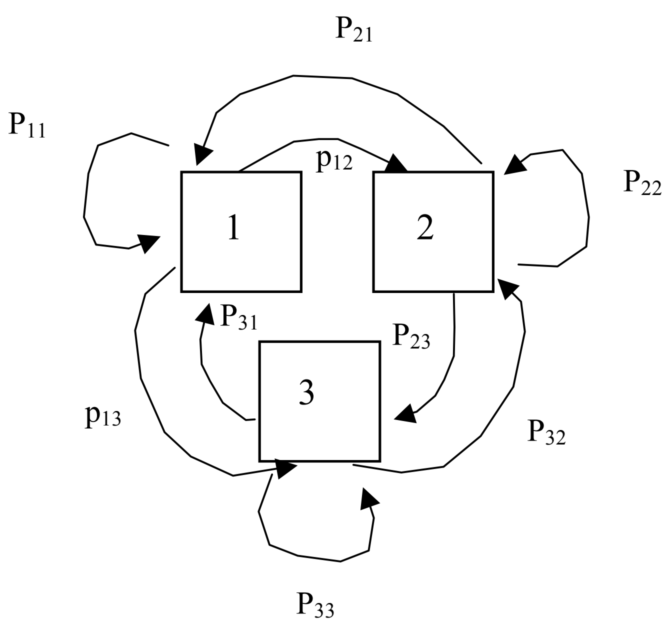
We can use a 2-state, first order Markov chain to:

a) create an artificial time series for, for example yes/no precipitation: From a precipitation series, we can estimate p_{00}, p_{11}, and therefore p_{01}, p_{10}. Then, if we are in state 0, we get a random number x between zero and 1. If $x \le p_{00}$, we stay in state 0, otherwise we go to state 1.

b) make a forecast: for example, given $x_t = 0$, we can predict $P(x_{t+1} = 0) = p_{00}$ and $P(x_{t+1} = 1) = p_{01}$ and so on.

We could also check for goodness of fit (Wilks, p104), comparing observed data histograms with simulated data with Markov chains.

Multistate first-order Markov chain



Again, the transition probabilities can be derived from the sample, and similar rules are valid, e.g., $p_{11} + p_{12} + p_{13} = 1$, etc.

Second order Markov chains:

$p_{ijk} = P(x_{t+1} = k \mid x_t = j, x_{t-1} = i)$ Becomes complicated!...