Trends:

Cyclical/Seasonal Trends

Week III: Video 9

STAT 485/685, Fall 2020, SFU

Sonja Isberg

Review: Trends in Time Series

Suppose our process of interest, $\{Y_t\}$, has some mean function μ_t (which may or may not be a function of t).

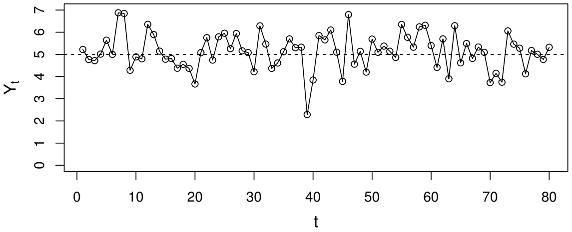
We can separate out the mean from the rest of the process by writing:

$$Y_t = \mu_t + X_t$$

where $\{X_t\}$ is the "de-trended" version of the process, i.e. $E(X_t) = 0$.

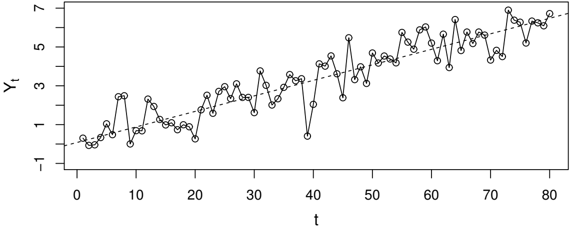
Some examples of how μ_t may look:

- Constant trend: $\mu_t = \mu$ for all t

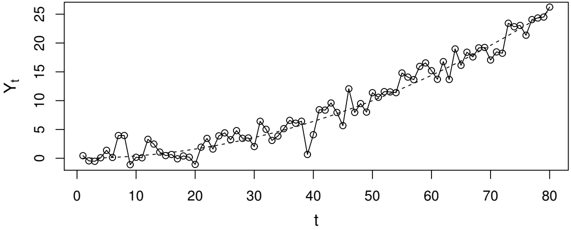


Review: Trends in Time Series (cont'd)

- Linear trend: $\mu_t = \beta_0 + \beta_1 t$

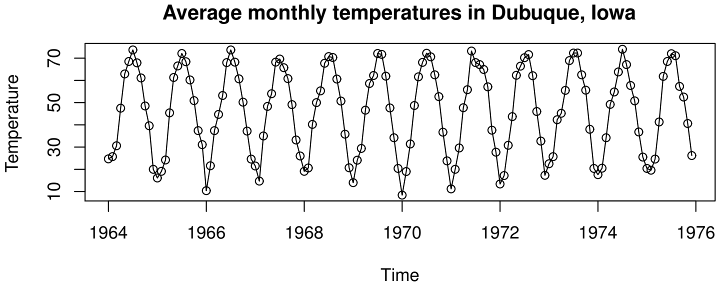


- Quadratic trend: $\mu_t = \beta_0 + \beta_1 t + \beta_2 t^2$



Review: Trends in Time Series (cont'd)

- Cyclical/seasonal trend: e.g., $\mu_t = \mu_{t-12}$ for all t



Video 9 Learning Objectives

By the end of this video, we should be able to:

- Recognize a seasonal trend from a given time series plot
- Obtain estimates of the mean at different times, using a *seasonal means model*
- Obtain estimates of the mean at different times, using a *cosine trend model*
- Identify some advantages and disadvantages of each of the two models

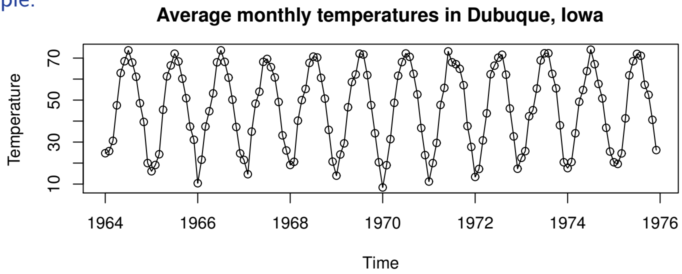
Cyclical or Seasonal Trend

Suppose:

$$Y_t = \mu_t + X_t,$$

where μ_t is some seasonal trend.

Example:



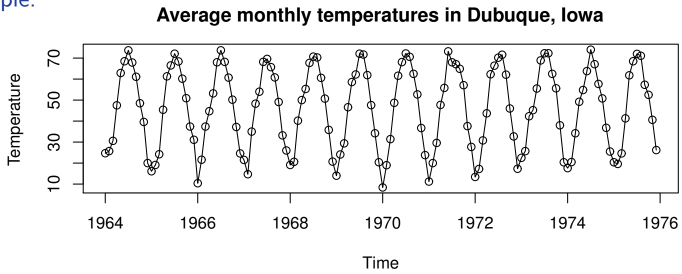
Cyclical or Seasonal Trend

Suppose:

$$Y_t = \mu_t + X_t,$$

where μ_t is some seasonal trend.

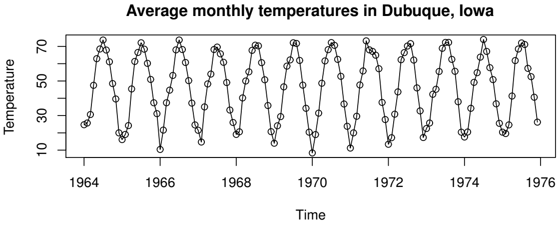
Example:



We will consider **two** possible ways of modelling μ_t:

- Seasonal means model
- Cosine trend model

Seasonal Means Model



We know that our data is seasonal by month, so we propose the model:

$$\mu_t = \begin{cases} \beta_1 & \text{for } t = 1, 13, 25, \ldots \\ \beta_2 & \text{for } t = 2, 14, 26, \ldots \\ \vdots & \\ \beta_{12} & \text{for } t = 12, 24, 36, \ldots \end{cases}$$

So, the parameter β_1 is the mean for January, β_2 is the mean for February, etc.

Seasonal Means Model: Estimating Parameters

$$\mu_t = \begin{cases} \beta_1 & \text{for } t = 1, 13, 25, \ldots \\ \beta_2 & \text{for } t = 2, 14, 26, \ldots \\ \vdots & \\ \beta_{12} & \text{for } t = 12, 24, 36, \ldots \end{cases}$$

Each parameter β_j is the mean for the j^{th} month (for $j = 1, 2, \ldots, 12$).

Each β_j can be estimated by taking the sample mean of the observations for that month:

$$\hat{\beta}_j = \frac{1}{N} \sum_{i=0}^{N-1} Y_{j+12i},$$

where N is the number of (complete) years of data.

Then, the estimate of the mean at month j is: $\hat{\mu}_j = \hat{\beta}_j$ (for $j = 1, 2, \ldots, 12$).

Seasonal Means Model: Estimating Parameters

$$\mu_t = \begin{cases} \beta_1 & \text{for } t = 1, 13, 25, \dots \\ \beta_2 & \text{for } t = 2, 14, 26, \dots \\ \vdots & \\ \beta_{12} & \text{for } t = 12, 24, 36, \dots \end{cases}$$

Each parameter β_j is the mean for the j^{th} month (for $j = 1, 2, \dots, 12$).

Each β_j can be estimated by taking the sample mean of the observations for that month:

$$\hat{\beta}_j = \frac{1}{N} \sum_{i=0}^{N-1} Y_{j+12i},$$

where N is the number of (complete) years of data.

Then, the estimate of the mean at month j is: $\hat{\mu}_j = \hat{\beta}_j$ (for $j = 1, 2, \dots, 12$).

Note: This $\hat{\beta}_j$ is similar to our estimate of the constant mean, $\hat{\mu} = \bar{Y}$, except we are only using every 12^{th} time point.

These computations will mostly be done by statistical software.

Seasonal Means Model: Properties of $\hat{\beta}_j$

Each β_j is estimated by taking the sample mean of the observations for that month.

This is similar to our estimate of the constant mean, $\hat{\mu} = \bar{Y}$, except we are only using every 12^{th} time point.

Seasonal Means Model: Properties of $\hat{\beta}_j$

Each β_j is estimated by taking the sample mean of the observations for that month.

This is similar to our estimate of the constant mean, $\hat{\mu} = \bar{Y}$, except we are only using every 12^{th} time point.

If we assume that $\{X_t\}$ is a stationary process with autocorrelation function ρ_k:

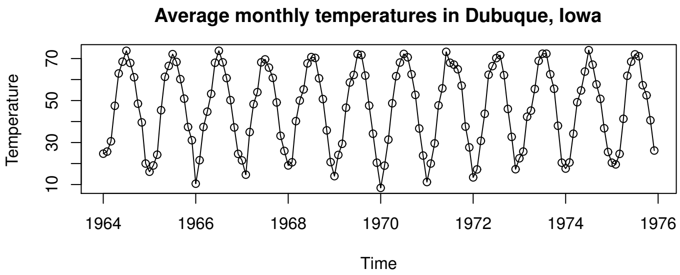
$$Var(\hat{\beta}_j) = \frac{\gamma_0}{N}\left[1 + 2\sum_{k=1}^{N-1}\left(1 - \frac{k}{N}\right)\rho_{12k}\right],$$

where N is the number of years of data.

Some special cases:

- If $\{X_t\}$ is a white noise process: $Var(\hat{\beta}_j) = \gamma_0/N$.
- Even if some of the ρ's are non-zero, as long as $\rho_{12k} = 0$ (i.e., observations between the same month are independent): $Var(\hat{\beta}_j) = \gamma_0/N$.

R Example



R Code:

```
data(tempdub)
month. <- season(tempdub)
model.seasonal <- lm(tempdub~month.-1)
summary(model.seasonal)
```

R Example(cont'd)

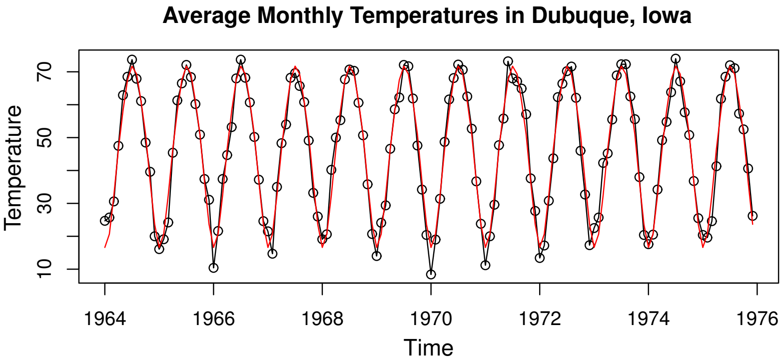
Output:

```
                Estimate Std. Error t value Pr(>|t|)    
month.January     16.608      0.987   16.83   <2e-16 ***
month.February    20.650      0.987   20.92   <2e-16 ***
month.March       32.475      0.987   32.90   <2e-16 ***
month.April       46.525      0.987   47.14   <2e-16 ***
month.May         58.092      0.987   58.86   <2e-16 ***
month.June        67.500      0.987   68.39   <2e-16 ***
month.July        71.717      0.987   72.66   <2e-16 ***
month.August      69.333      0.987   70.25   <2e-16 ***
month.September   61.025      0.987   61.83   <2e-16 ***
month.October     50.975      0.987   51.65   <2e-16 ***
month.November    36.650      0.987   37.13   <2e-16 ***
month.December    23.642      0.987   23.95   <2e-16 ***
```

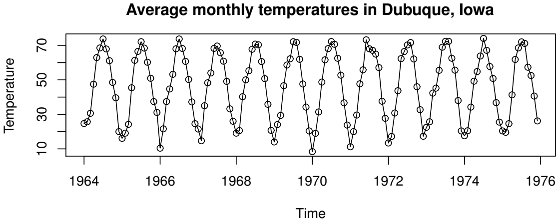
R Example (cont'd)

Plotting:

```
plot(tempdub, type='o', xlab='Time', ylab='Temperature',
     main='Average Monthly Temperatures in Dubuque, Iowa')
fitted.vals <- ts(fitted(model.seasonal), freq=12, start=c(1964,1))
lines(fitted.vals, col='red')
```



Cosine Trend Model



Some drawbacks of the seasonal means model:

- Requires the estimation of 12 parameters.
- Parameters are considered to be independent, and therefore do not take the shape of the trend into account.

A **cosine trend model** can incorporate the smooth change expected from one time point to the next.

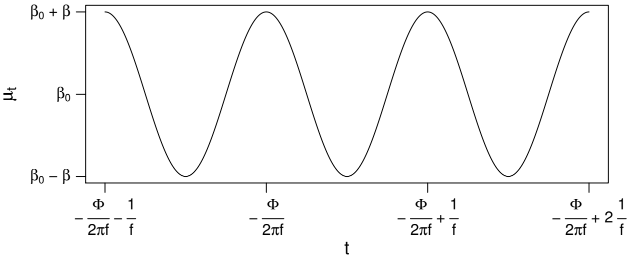
Cosine Trend Model (cont'd)

Suppose:

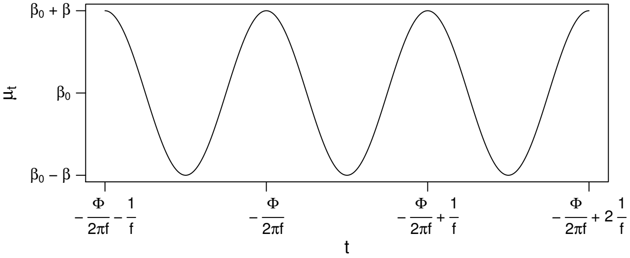
$$\mu_t = \beta\, cos\left[2\pi f\left(t + \frac{\Phi}{2\pi f}\right)\right] + \beta_0.$$

This model contains the parameters:

- β (> 0): *amplitude* of the curve
- f: *frequency* of the curve
- Φ: *phase* of the curve
- β_0: vertical shift of the curve



Cosine Trend Model (cont'd)



What we know:

- The cycle repeats every $1/f$ time units. So, for monthly seasonal data we set $f = 1/12$.
- Oscillates between a maximum of $\beta_0 + \beta$ and a minimum of $\beta_0 - \beta$.
- Φ serves to set the time at which the curve starts.

Cosine Trend Model: Estimating Parameters

The model can be re-written as:

$$\mu_t = \beta_0 + \beta_1\, cos(2\pi ft) + \beta_2\, sin(2\pi ft)$$

We have only three unknown parameters: β_0, β_1 and β_2.

The parameters can be estimated using a linear regression, where instead of t our predictors are $cos(2\pi ft)$ and $sin(2\pi ft)$. These computations will mostly be done by statistical software.

Then, the estimate of the mean at time t is:

$\hat{\mu}_t = \hat{\beta}_0 + \hat{\beta}_1\, cos(2\pi ft) + \hat{\beta}_2\, sin(2\pi ft)$.

Cosine Trend Model: Estimating Parameters

The model can be re-written as:

$$\mu_t = \beta_0 + \beta_1 \, cos(2\pi ft) + \beta_2 \, sin(2\pi ft)$$

We have only three unknown parameters: β_0, β_1 and β_2.

The parameters can be estimated using a linear regression, where instead of t our predictors are $cos(2\pi ft)$ and $sin(2\pi ft)$. These computations will mostly be done by statistical software.

Then, the estimate of the mean at time t is:
$\hat{\mu}_t = \hat{\beta}_0 + \hat{\beta}_1 \, cos(2\pi ft) + \hat{\beta}_2 \, sin(2\pi ft)$.

Caution: We must be very careful when defining f for monthly seasonality. If our data records the time points as $t = 1$ for January, $t = 2$ for February, etc., then the monthly pattern repeats itself every 12 time points, and we should be using $f = 1/12$.
If, however, our data records the time points as each time point being one *year*, then the monthly pattern repeats itself every 1 time point, and we should be using $f = 1$.

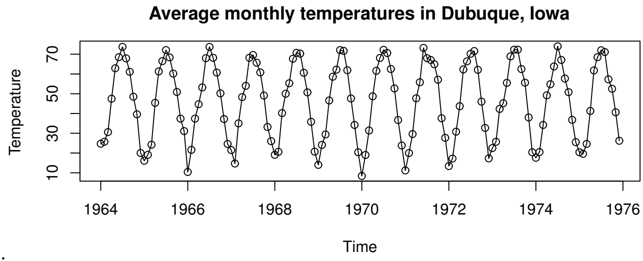
Cosine Trend Model: Properties of $\hat{\beta}_0$, $\hat{\beta}_1$ and $\hat{\beta}_2$

The estimates of β_0, β_1 and β_2 are obtained using linear regression, where instead of t our predictors are $cos(2\pi ft)$ and $sin(2\pi ft)$.

Analytical expressions exist for $Var(\hat{\beta}_0)$, $Var(\hat{\beta}_1)$ and $Var(\hat{\beta}_2)$, but they are quite complicated.

If $\{X_t\}$ is a white noise process: $Var(\hat{\beta}_1) = Var(\hat{\beta}_2) = 2\gamma_0/n$.

R Example



R Code:

```
data(tempdub)
har. <- harmonic(tempdub, 1)
model.seasonal2 <- lm(tempdub~har.)
summary(model.seasonal2)
```

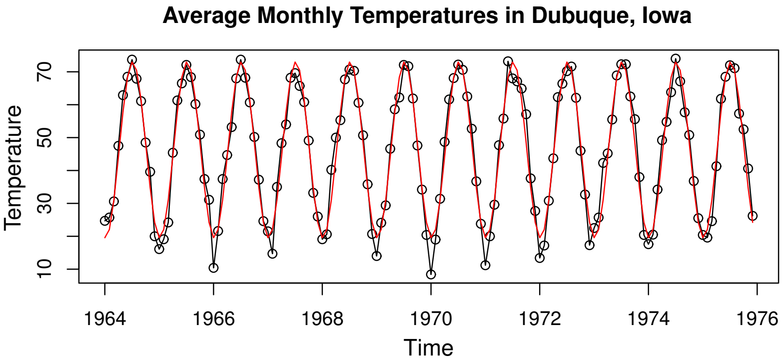
Output:

```
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)       46.2660     0.3088 149.816  < 2e-16 ***
har.cos(2*pi*t) -26.7079     0.4367 -61.154  < 2e-16 ***
har.sin(2*pi*t)  -2.1697     0.4367  -4.968 1.93e-06 ***
```

R Example (cont'd)

Plotting:

```
plot(tempdub, type='o', xlab='Time', ylab='Temperature',
     main='Average Monthly Temperatures in Dubuque, Iowa')
fitted.vals2 <- ts(fitted(model.seasonal2), freq=12, start=c(1964,1))
lines(fitted.vals2, col='red')
```



Seasonal Means Model vs. Cosine Trend Model

Both the seasonal means model and cosine trend model are fairly easy to fit using R. They are both fitted using least-squares estimates.

The cosine trend model:

- Requires the estimation of fewer parameters.
- Takes the shape of the trend into account.
- However: Has a strong model assumption.

Seasonal Means Model vs. Cosine Trend Model

Both the seasonal means model and cosine trend model are fairly easy to fit using R. They are both fitted using least-squares estimates.

The cosine trend model:

- Requires the estimation of fewer parameters.
- Takes the shape of the trend into account.
- However: Has a strong model assumption.

Assume $\{X_t\}$ is a white noise process. Let's compare the estimates of the mean in January:

- Seasonal means model: $\hat{\mu}_1 = \hat{\beta}_1$. Then: $Var(\hat{\mu}_1) = Var(\hat{\beta}_1) = \gamma_0/N$.
- Cosine trend model: $\hat{\mu}_1 = \hat{\beta}_0 + \hat{\beta}_1 \cos\left(\frac{2\pi}{12}\right) + \hat{\beta}_2 \sin\left(\frac{2\pi}{12}\right)$. Then: $Var(\hat{\mu}_1) = \cdots = 3\gamma_0/n$.

Seasonal Means Model vs. Cosine Trend Model

Both the seasonal means model and cosine trend model are fairly easy to fit using R. They are both fitted using least-squares estimates.

The cosine trend model:

- Requires the estimation of fewer parameters.
- Takes the shape of the trend into account.
- However: Has a strong model assumption.

Assume $\{X_t\}$ is a white noise process. Let's compare the estimates of the mean in January:

- Seasonal means model: $\hat{\mu}_1 = \hat{\beta}_1$. Then: $Var(\hat{\mu}_1) = Var(\hat{\beta}_1) = \gamma_0/N$.
- Cosine trend model: $\hat{\mu}_1 = \hat{\beta}_0 + \hat{\beta}_1 \, cos\left(\frac{2\pi}{12}\right) + \hat{\beta}_2 \, sin\left(\frac{2\pi}{12}\right)$. Then: $Var(\hat{\mu}_1) = \cdots = 3\gamma_0/n$.
- Since generally $N \approx n/12$:

$$\frac{\text{Variance of } \hat{\mu}_1 \text{ for seasonal means model}}{\text{Variance of } \hat{\mu}_1 \text{ for cosine trend model}} = \frac{\gamma_0/(n/12)}{3\gamma_0/n} = 4$$

Final Comments

That's all for now!

In this video, we've learned about two different approaches to modelling a seasonal trend: the **seasonal means model** and the **cosine trend model.**

We've also learned about some of the properties of the estimates arising from each of these models.

Coming Up Next: Residual analysis, and determining whether a chosen model matches the data well.

References

[1] Cryer, J. D., & Chan, K. S. (2008). *Time series analysis: with applications in R.* Springer Science and Business Media.

[2] Chan, K. S., & Ripley, B. (2020). TSA: Time Series Analysis. R package version 1.2.1.