STAT 400 Data Analysis Sept. 30, 2005

Today:

Applying Seasonal Adjustment (based on Weekly data) to a day-of-the-week.
More about Time Series Fitting (pp 152-158)

Seasonal Adjustment: Use

annual mean + (original data – loess fit) [additive]

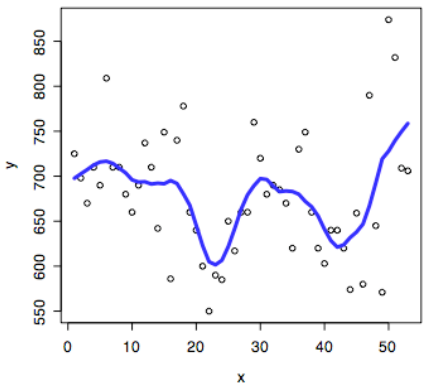
or

annual mean * (original data/loess fit) [multiplicative]

1. Additive method

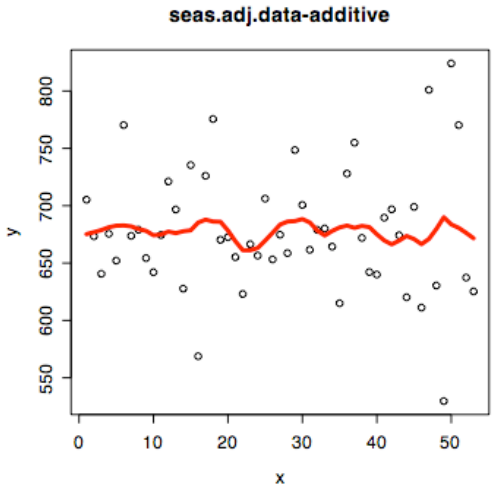
```
> smo
function (x,y,span=.5,col="blue",main=main,...)
{
yl=loess(y~x,span=span,degree=1)
plot(x,y,col="black",type="p",main=main)
smooth=predict(yl)
lines(x,smooth,col=col,lwd=3)
invisible(smooth=smooth)
}

smooth=smo(week,WkD,span=.2)
```

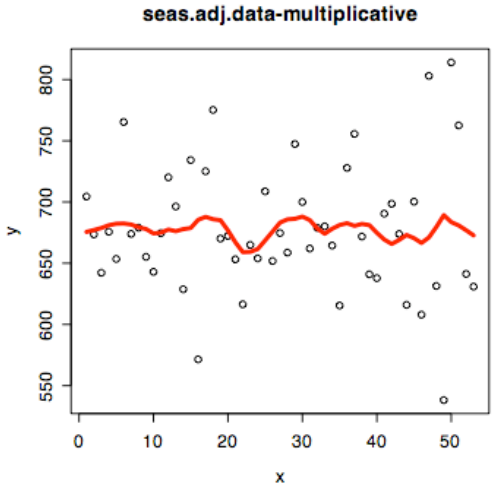



```
mean(WkD)
[1] 677.9811

seas.adj.D=678+(WkD-smooth)
> smo(week,seas.adj.D,col="red",span=.2,main="seas.adj.data-additive")
```



seas.adj.D=678*(WkD/smooth)
> smo(week,seas.adj.D,col="red",span=.2,main="seas.adj.data-multiplicative")



The result is almost the same (but not exactly). The advantage of the multiplicative version is that the vector of ratios (WkD/678) can be applied multiplicatively to the vector of data for single day-of-the-week. (Both Deliveries and Sales).

Time Series Fitting

See p 152: melanoma data

> attach(melanoma.df)
> colnames(melanoma.df)
[1] "year" "incidence"
> plot(year,incidence,type="b",main="Melanoma Incidence,1936-1972")

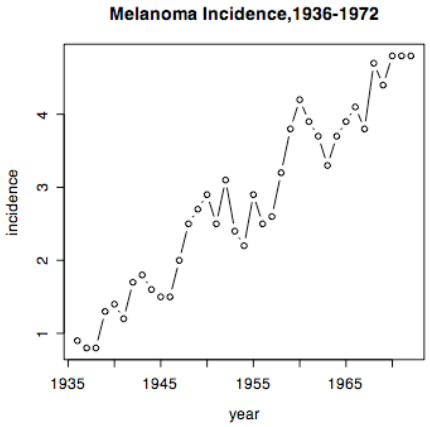


Fig 3.65 is loess! (not a straight line)

smo(year,incidence,main="loess fit - melanoma data",type="n",span=.75)

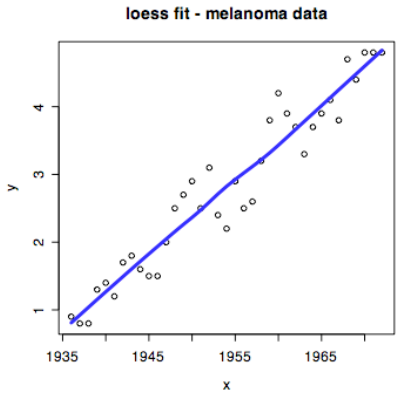
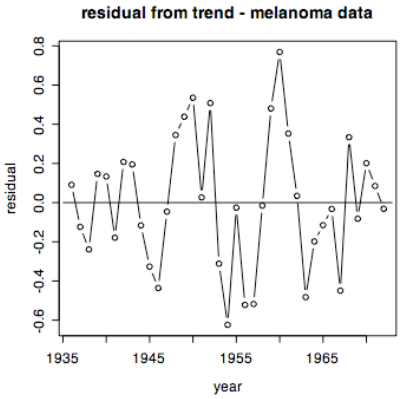


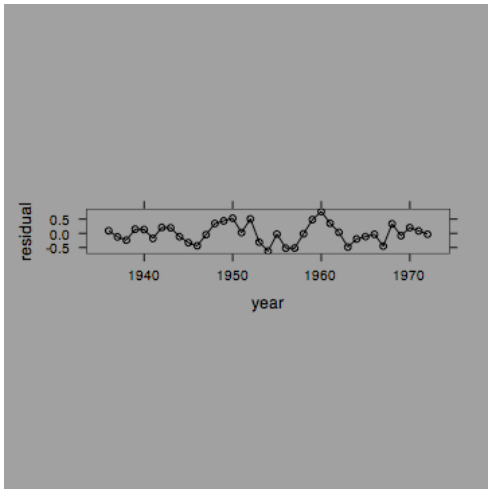
Fig 3.66 is resids

> plot(year,residual,main="residual from trend - melanoma data",type="b")
> lines(c(1935,1973),c(0,0))



Cleveland likes the banked version:

xyplot(residual~year,aspect="xy",type="b",col="black")



Hard to see though ...
a=smo(year,incidence,span=.75)
> resids=incidence-a
> smo(year,resids,span=.25)

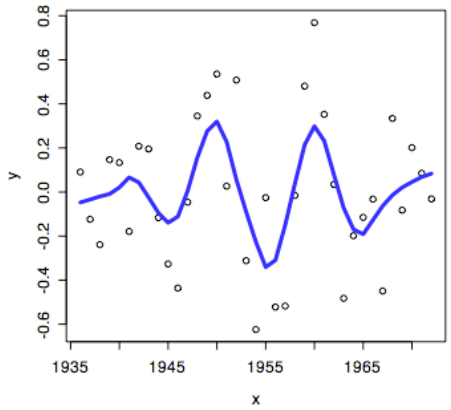
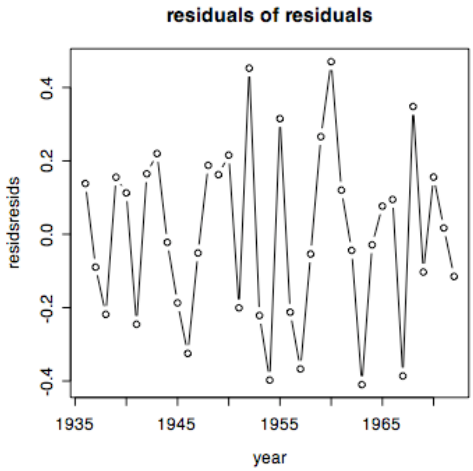


Fig 3.67 is loess of resids

Fig 3.68 is resids from 3.67

b=smo(year,resids,span=.25)
> residsresids=resids-b
> plot(year,residsresids,main="residuals of residuals",type="b")

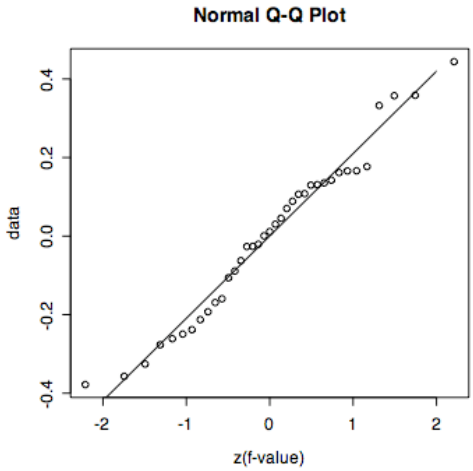


Summary Fig 3.70 p 157

p 158 shows smoothed residuals and sunspot cycle – useful theory! (Wear a hat)

p 156 Cleveland checks normality of residuals. Here is a program for normal q-q plots.

```
> norm.qq.plot
function (x)
{
        n=length(x)
        m=mean(x)
        s=sd(x)
        f=(1:n-0.5)/n
        x=sort(x)
        q=qnorm(f)
        plot(q,x,xlab="z(f-value)",ylab="data",main="Normal Q-Q Plot")
        lines(c(-2,2),c(qnorm(.023,m,s),qnorm(.977,m,s)))
}
```



Note re time series: fit more than one component, since each component may have a different cause.

Next time: Rest of Ch 3 and start of Ch 4.
Homework for Monday Oct 3: Prepare discussion of pp 159-179

Seasonal loess: Rough periodic (like melanoma incidence or sunspots) is not like seasonal.

p 161 Use loess for sub-cycle (e.g. 32 Januarys) and then combine. Result like p 163.

Cycle plots – another way to display seasonal data.