

University of West Florida

Time Series Analysis STA 6856 | Professor Dr. Tharindu De Alwis

Edwin Quijano

Friday, September 26, 2025

1. (20 points) Consider the MA(1) model $Y_t = -\theta e_{t-1} + e_t$ where $e_t \sim WN(0, \sigma^2)$

(a) Find $E(Y_t)$

$$Y_t = -\theta e_{t-1} + e_t \text{ where } e_t \sim WN(0, \sigma^2)$$

$$E(Y_t) = 0 \text{ for all } t$$

$$Var(Y_t) = \theta^2 \text{ for all } t$$

$$Cov(Y_t, Y_t)$$

$$Y_t = \mu - \theta e_{t-1} + e_t$$

$$E(Y_t) = E(\mu) - \theta E(e_{t-1}) + E(e_t)$$

$$E(\mu) = \mu \text{ Constant}$$

$$E(e_{t-1}) = 0 \text{ W N}$$

$$E(e_t) = 0 \text{ W N}$$

$$E(Y_t) = \mu - \theta(0) + 0 = \mu$$

(b) Find $V(Y_t) = Cov(Y_t, Y_t)$

$Y_t = -\theta e_{t-1} + e_t$ where e_t is white noise

$$V(Y_t) - \mu = -\theta e_{t-1} + e_t$$

$$Var(Y_t) = Var(-\theta e_{t-1} + e_1)$$

$$V(Y_t) = Var(-\theta e_{t-1} + e_t)$$

$$= \theta^2 Var(e_{t-1}) + Var(e_t) + 2(-\theta(1) Cov(e_{t-1} + e_t))$$

$$= \theta^2 \sigma^2 + \sigma^2 + 0 = (1 + \theta^2) \sigma^2$$

(c) Find $Cov(Y_t, Y_{t-1})$

$$Y_t - \mu = -\theta e_{t-1} + e_t$$

$$Y_{t-1} - \mu = -\theta e_t - 2 + e_{t-1}$$

$$\begin{aligned} Cov(Y_t, Y_{t-1}) &= Cov(-\theta e_{t-1} + e_t, -\theta e_{t-2} + e_{t-1}) \\ &= Cov(-\theta e_{t-1}, -\theta e_{t-2}) + Cov(-\theta e_{t-1}, e_{t-1}) + Cov(e_t, -\theta e_{t-2}) + Cov(e_t, e_{t-1}) \\ &= Cov(-\theta e_{t-1}, -\theta e_{t-2}) = 0 \\ &+ Cov(-\theta e_{t-1}, e_{t-1}) = \text{nonzero} \\ &+ Cov(e_t, -\theta e_{t-2}) = 0 \\ &+ Cov(e_t, e_{t-1}) = 0 \end{aligned}$$

$$(t-1, t-2) = 0$$

$$(t, t-2) = 0$$

$$(t, t-1) = 0$$

$$Cov(-\theta e_{t-1}, e_{t-1}) = \theta Var(e_{t-1}) = \theta \sigma^2$$

$$Cov(Y_t, Y_{t-1}) = \theta \sigma^2$$

(d) It can be shown that $Cov(Y_t, Y_{t-k}) = 0$ for $k > 1$. Is Y_t stationary? Explain briefly.

$k > 1$ has no shared noise terms in $Y_t - \mu = -\theta e_{t-1} + e_t$ and $Y_{t-k} = -\theta e_{t-k-1} + e_{t-k}$

$$Cov(Y_t, Y_{t-k}) = 0, k > 1$$

Is Weakly stationary since Covariance depends only on the lag K .

2. (10 points) Consider the MA(3) model with $\theta_1 = 0.8$, $\theta_2 = 0.6$, and $\theta_3 = 0.4$.

(a) Find the theoretical ACF.

```
library(stats)
```

```
ARMAacf(ma=0.8, lag.max=10)
```

	0	1	2	3	4	5	6	7
1.0000000	0.4878049	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000
	8	9	10					
0.0000000	0.0000000	0.0000000						

```
ARMAacf(ma=0.6, lag.max=10)
```

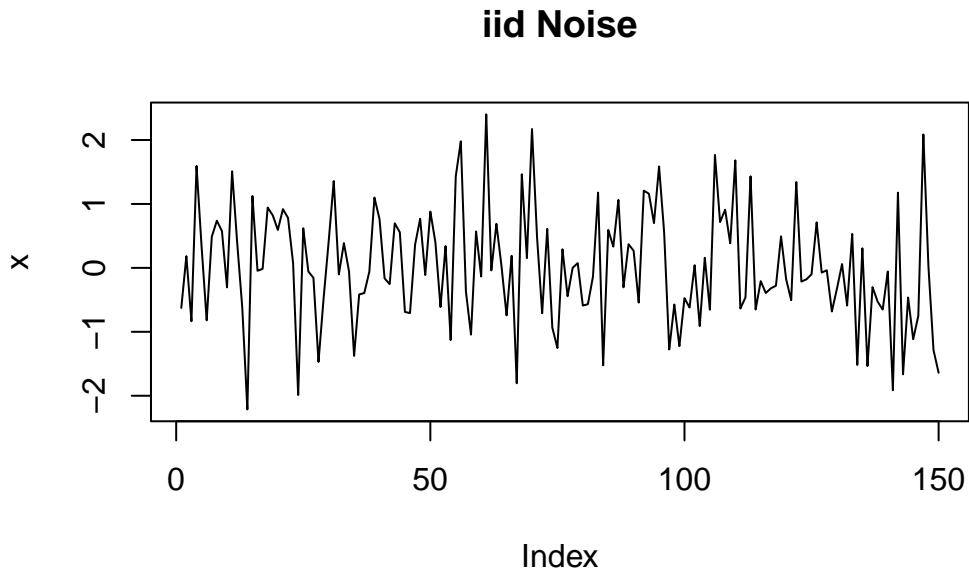
	0	1	2	3	4	5	6	7
1.0000000	0.4411765	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000
	8	9	10					
0.0000000	0.0000000	0.0000000						

```
ARMAacf(ma=0.4, lag.max=10)
```

	0	1	2	3	4	5	6	7
1.0000000	0.3448276	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000
	8	9	10					
0.0000000	0.0000000	0.0000000						

(b) Generate $n = 150$ observations of this MA(3) time series. Then, plot the sample autocorrelation function (sample ACF).

```
set.seed(1) #set.seed will generate the same random sample again and again
x=rnorm(150)
plot(x,main="iid Noise",type="l")
```



3. (50 points) The data file `winnebago` contains monthly unit sales of recreational vehicles from Winnebago, Inc., from November 1966 through February 1972.

Use the following pair of codes to import the dataset

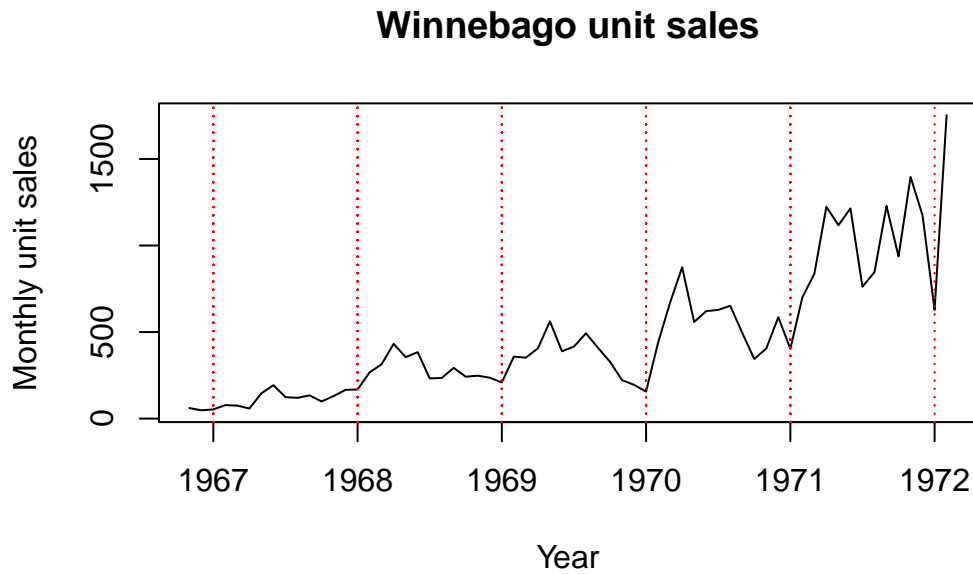
```
library(TSA)

data("winnebago")

y<-winnebago
t<-time(y)
```

(a) Display and interpret the time series plot for these data.

```
plot(y, ylab="Monthly unit sales", xlab="Year",  
     main="Winnebago unit sales")  
abline(v=floor(time(y)), col="red", lty=3)
```



Sales rise sharply from late-1966 to early-1972 — a clear upward trend.

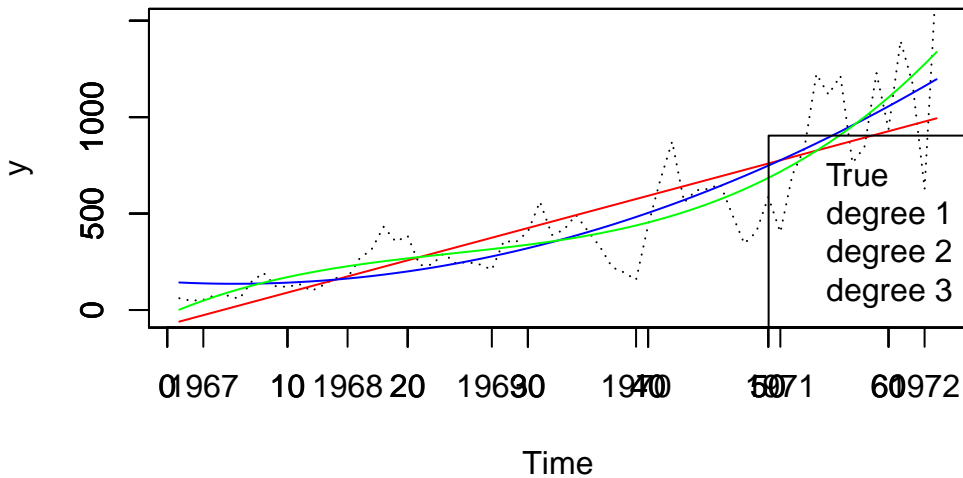
Each year, sales peak in late spring/summer and dip in winter, showing strong 12-month seasonality.

The seasonal ups and downs get larger as sales grow, so variability increases with the level that's heteroskedasticity and suggests a multiplicative seasonal pattern.

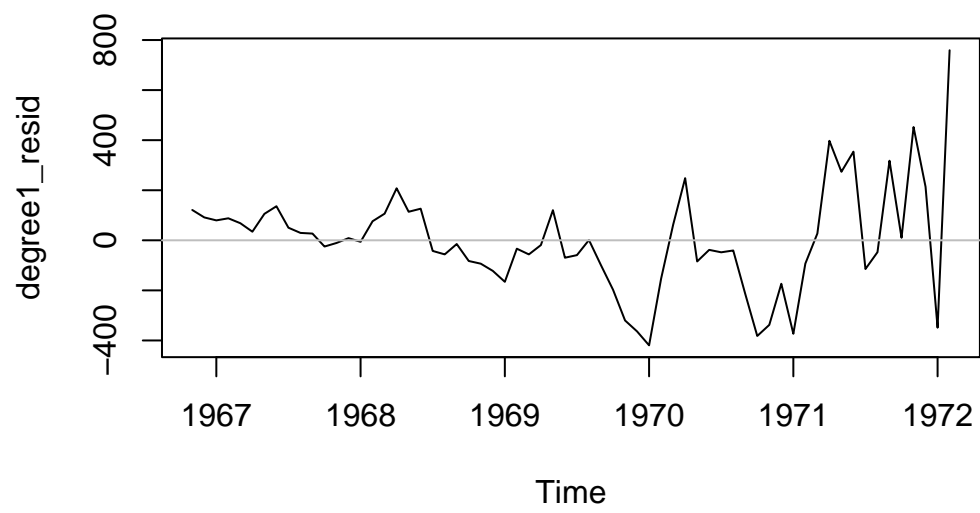
(b) Use appropriate method to fit the trend in this data. Interpret the regression output. Plot the standardized residuals from the fit as a time series. Interpret the plot.

```
reg.trend1 = trend(y, 1) #linear function
reg.trend2 = trend(y, 2) #Quadratic function
reg.trend3 = trend(y,3) # Cubic function

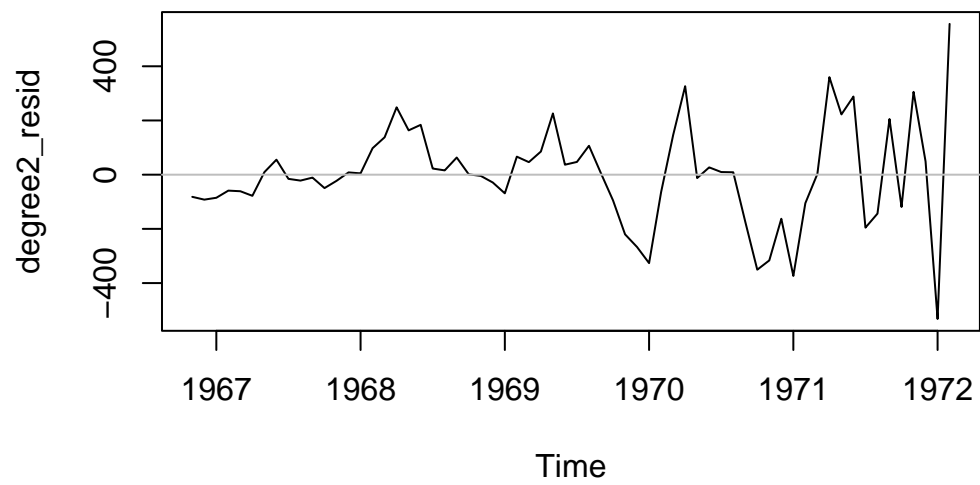
plot(y,type="l",ylim=c(-30,1500),lty=3)
par(new=TRUE)
plot(reg.trend1,col="red",ylim=c(-30,1500),type="l",lwd=1,ylab="",xlab="")
par(new=TRUE)
plot(reg.trend2,col="blue",ylim=c(-30,1500),type="l",lwd=1,ylab="",xlab="")
par(new=TRUE)
plot(reg.trend3,col="green",ylim=c(-30,1500),type="l",lwd=1,ylab="",xlab="")
legend("bottomright",legend=c("True","degree 1","degree 2","degree 3"),col=c("black ","red",
```



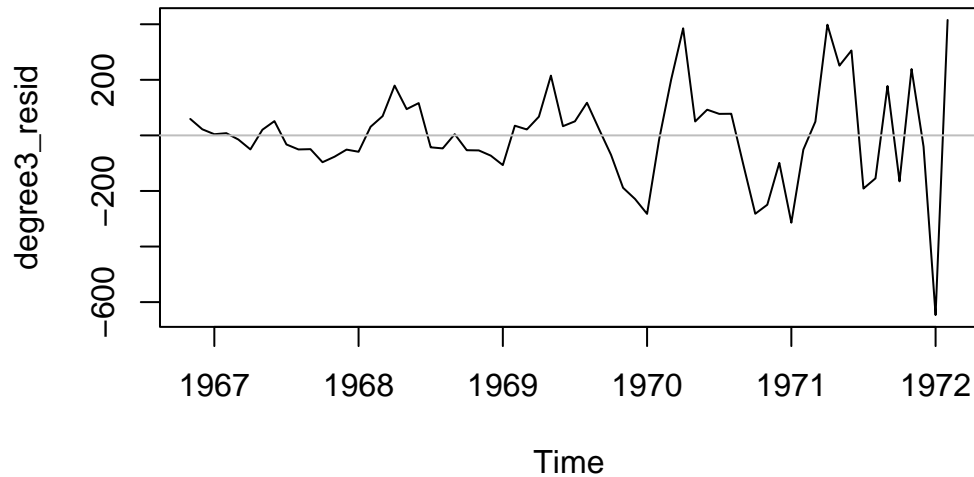
```
degree1_resid=y-reg.trend1
degree2_resid=y-reg.trend2
degree3_resid=y-reg.trend3
plot(degree1_resid)
abline(h=0, col="gray")
```



```
plot(degree2_resid)  
abline(h=0, col="gray")
```




```
plot(degree3_resid)
abline(h=0, col="gray")
```



```
fit_quad <- lm(as.numeric(y) ~ t + I(t^2))
summary(fit_quad)
```

Call:

```
lm(formula = as.numeric(y) ~ t + I(t^2))
```

Residuals:

Min	1Q	Median	3Q	Max
-532.59	-82.60	3.66	71.18	556.36

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	1.736e+08	4.295e+07	4.043	0.000150	***
t	-1.765e+05	4.361e+04	-4.048	0.000148	***
I(t^2)	4.487e+01	1.107e+01	4.052	0.000146	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

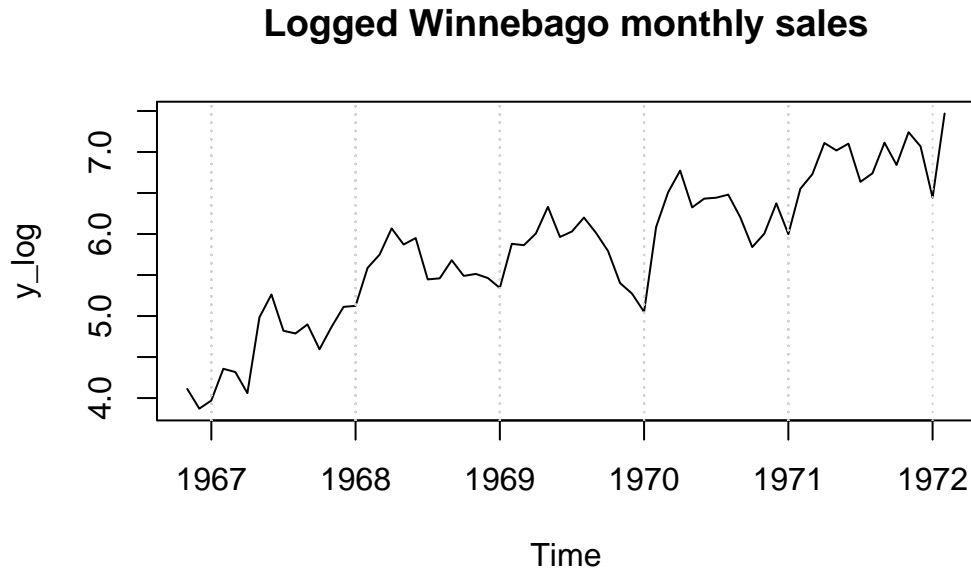
Residual standard error: 187.7 on 61 degrees of freedom
Multiple R-squared: 0.7569, Adjusted R-squared: 0.7489
F-statistic: 94.97 on 2 and 61 DF, p-value: $< 2.2e-16$

The standardized residuals still show a clear seasonal pattern positive in summer, negative in winter, and month-to-month correlation. They're not behaving like random noise.

Residual spread is larger later in time, so the variance isn't constant.

(c) Now take natural logarithms of the monthly sales figures and display and interpret the time series plot of the transformed values.

```
y_log = log(y)
plot.ts(y_log,
main="Logged Winnebago monthly sales")
abline(v=floor(time(y)), col="lightgray", lty=3)
```



Trend remains upward, but seasonal amplitude becomes roughly constant across time.

The log makes the process look closer to additive (trend + seasonality + noise) with more stable variance.

(d) Use appropriate method to fit the trend in this logged data. Display and interpret the time series plot of the standardized residuals from this fit.

```
fit_log_lin <- lm(as.numeric(y_log) ~ t)
fit_log_quad <- lm(as.numeric(y_log) ~ t + I(t^2))

AIC(fit_log_lin, fit_log_quad)
```

```
              df      AIC
fit_log_lin   3 66.34478
fit_log_quad  4 64.55703
```

```
summary(fit_log_lin)
```

Call:

```
lm(formula = as.numeric(y_log) ~ t)
```

Residuals:

Min	1Q	Median	3Q	Max
-1.03669	-0.20823	0.04995	0.25662	0.86223

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-984.93878	62.99472	-15.63	<2e-16 ***
t	0.50306	0.03199	15.73	<2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

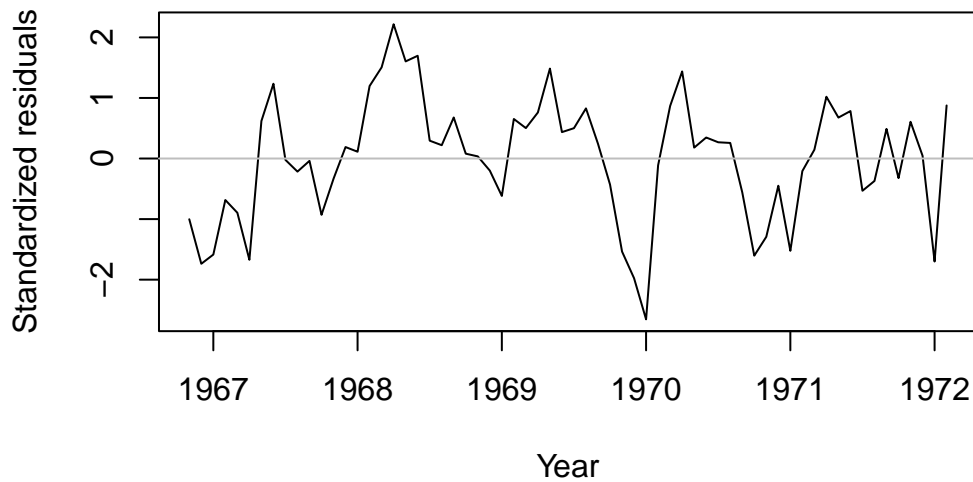
Residual standard error: 0.3939 on 62 degrees of freedom

Multiple R-squared: 0.7996, Adjusted R-squared: 0.7964

F-statistic: 247.4 on 1 and 62 DF, p-value: < 2.2e-16

```
r_std_log <- rstandard(fit_log_lin)
plot(ts(r_std_log, start=start(y), frequency=frequency(y)),
     ylab="Standardized residuals", xlab="Year",
     main="Std. residuals from linear trend (log scale)")
abline(h=0, col="gray")
```

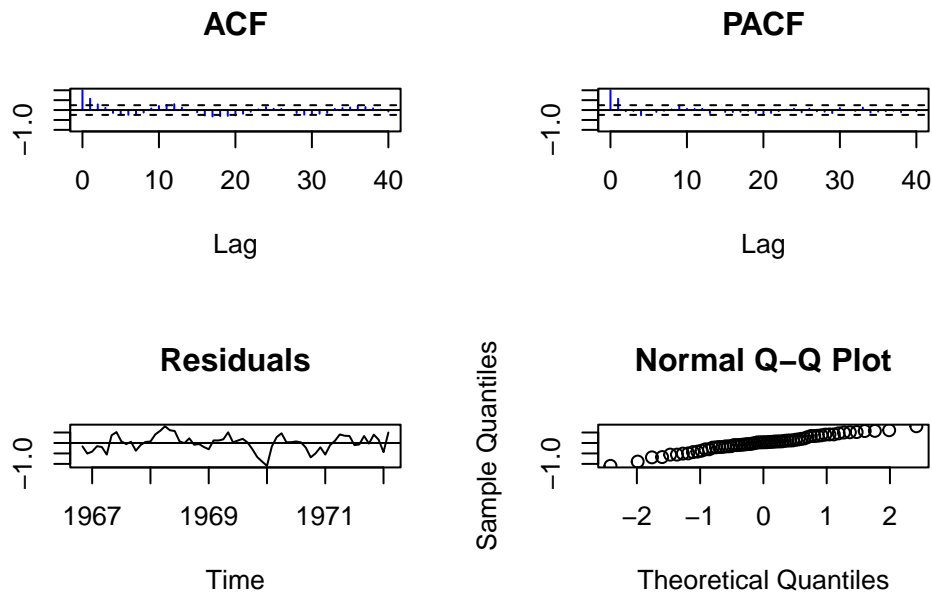
Std. residuals from linear trend (log scale)



```
# Estimation
wnbg.data.model = c("log","trend",2,"season",4)
r.wnbg = Resid(y,wnbg.data.model)
test(r.wnbg)
```

Null hypothesis: Residuals are iid noise.

Test	Distribution	Statistic	p-value
Ljung-Box Q	$Q \sim \text{chisq}(20)$	101.25	0 *
McLeod-Li Q	$Q \sim \text{chisq}(20)$	36.05	0.0152 *
Turning points T	$(T-41.3)/3.3 \sim N(0,1)$	31	0.0019 *
Diff signs S	$(S-31.5)/2.3 \sim N(0,1)$	32	0.8299
Rank P	$(P-1008)/86.3 \sim N(0,1)$	1033	0.7721



The time term is significant, sales are rising at a fairly steady percentage rate.

The residuals have roughly constant spread now (variance stabilized), which is better than on the raw scale.

However, the residual plot still shows seasonal waves and month-to-month correlation (big spikes at lags 1–2 and 12 in the ACF),

(e) Perform a runs test and Shapiro-Wilk test on the standardized residuals and interpret the results. [hint: shapiro.test(x)]

Recall: The Shapiro-Wilk's test or Shapiro test is a normality test in statistics. The null hypothesis of Shapiro's test is that the population is distributed normally.

```
shap_res <- shapiro.test(r.wnbg) # H0: normality
shap_res
```

Shapiro-Wilk normality test

```
data:  r.wnbg
W = 0.97996, p-value = 0.3832
```

H_0) : residuals are normally distributed.

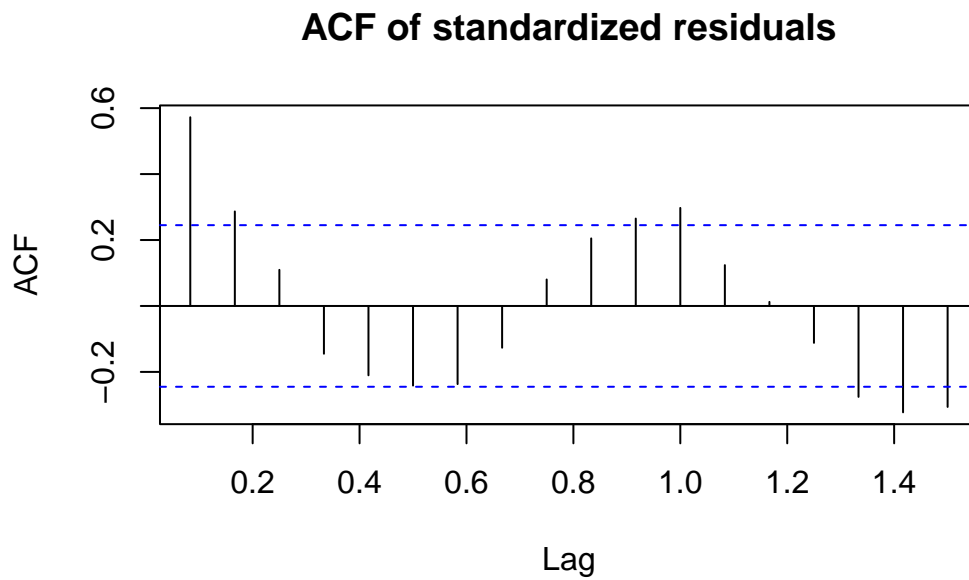
Interpretation: With $p = 0.38$ (> 0.05), fail to reject normality. The residuals are consistent with a normal distribution; W close to 1 also supports this.

Strong evidence the residuals are not random/independent. This means the model hasn't captured all the time-based structure (ej. seasonality/AR effects).

Shapiro-Wilk test: No evidence against normality of residuals. Their shape is roughly bell-shaped.

(f) Calculate and interpret the sample autocorrelations for the standardized residuals.

```
acf(r.wnbg, main="ACF of standardized residuals")
```



The ACF shows big spikes at lag 1 and lag 2, meaning residuals in back-to-back months still move together short-term persistence.

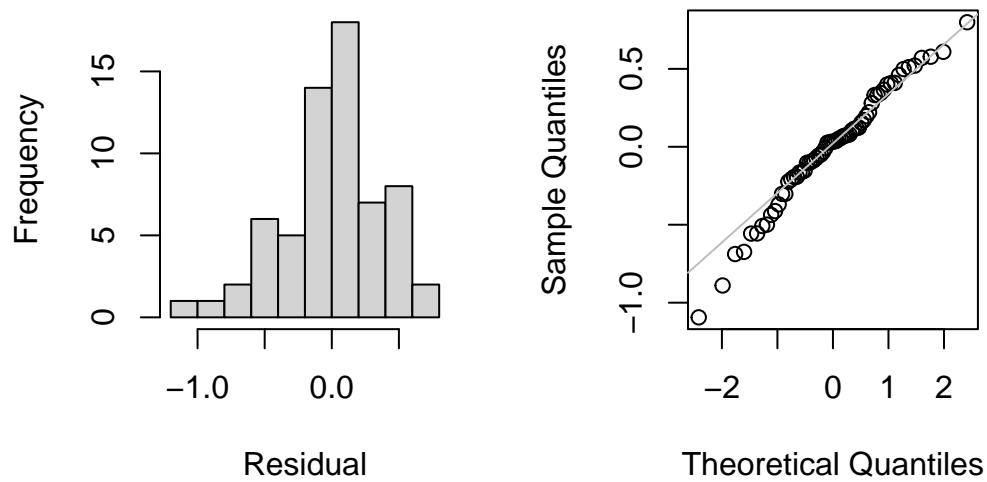
There's also a large spike at lag 12, signaling a yearly seasonal cycle left in the residuals.

residuals are not white noise.

(g) Investigate the normality of the standardized residuals (error terms). Consider histograms and normal probability plots. Interpret the plots.

```
op <- par(mfrow=c(1,2))
hist(r.wnbg, breaks="FD", main="Histogram of std. residuals",
      xlab="Residual", freq=TRUE)
qqnorm(r.wnbg, main="Normal Q-Q plot of std. residuals")
qqline(r.wnbg, col="gray")
```


Histogram of std. residual: Normal Q-Q plot of std. resid



```
par(op)
```

Roughly bell-shaped with no glaring multi-modality. There's mild tail heft but nothing major.

Residuals look approximately normal. Any minor tail issues are far less important than the leftover seasonality/autocorrelation seen earlier.