

sta457A2

2025-05-22

1. Linear trend with additive noise

Let

$$x_t = \beta_0 + \beta_1 t + w_t, \quad \mathbb{E}[w_t] = 0, \quad \text{Var}(w_t) = \sigma_w^2, \quad \{w_t\} \text{ i.i.d.}$$

%_____

(a) Non-stationarity of x_t Mean

$$\mu_x(t) = \mathbb{E}[x_t] = \beta_0 + \beta_1 t.$$

Variance

$$\text{Var}(x_t) = \text{Var}(\beta_0 + \beta_1 t + w_t) = \text{Var}(w_t) = \sigma_w^2.$$

Although the variance is constant, the mean depends on t ; therefore the process is not weakly stationary.

%_____

(b) Stationarity of the first difference $\nabla x_t := x_t - x_{t-1}$ Expression

$$\nabla x_t = (\beta_0 + \beta_1 t + w_t) - (\beta_0 + \beta_1(t-1) + w_{t-1}) = \beta_1 + (w_t - w_{t-1}).$$

Mean

$$\mathbb{E}[\nabla x_t] = \beta_1 + \mathbb{E}[w_t] - \mathbb{E}[w_{t-1}] = \beta_1.$$

Variance

$$\text{Var}(\nabla x_t) = \text{Var}(w_t - w_{t-1}) = \text{Var}(w_t) + \text{Var}(w_{t-1}) - 2 \text{Cov}(w_t, w_{t-1}) = 2\sigma_w^2.$$

Autocovariance

$$\gamma_{\nabla x}(h) = \text{Cov}(\nabla x_t, \nabla x_{t+h}) = \text{Cov}(w_t - w_{t-1}, w_{t+h} - w_{t+h-1}).$$

Because the w_t are independent,

$$\gamma_{\nabla x}(h) = \begin{cases} 2\sigma_w^2, & h = 0, \\ -\sigma_w^2, & h = \pm 1, \\ 0, & |h| \geq 2. \end{cases}$$

Both the mean and the autocovariance depend only on the lag h , not on calendar time, so $\{\nabla x_t\}$ is weakly stationary.

```
#####
## Question 1 - Parts (c) and (d) ##
## Linear-trend-plus-noise process ##
#####

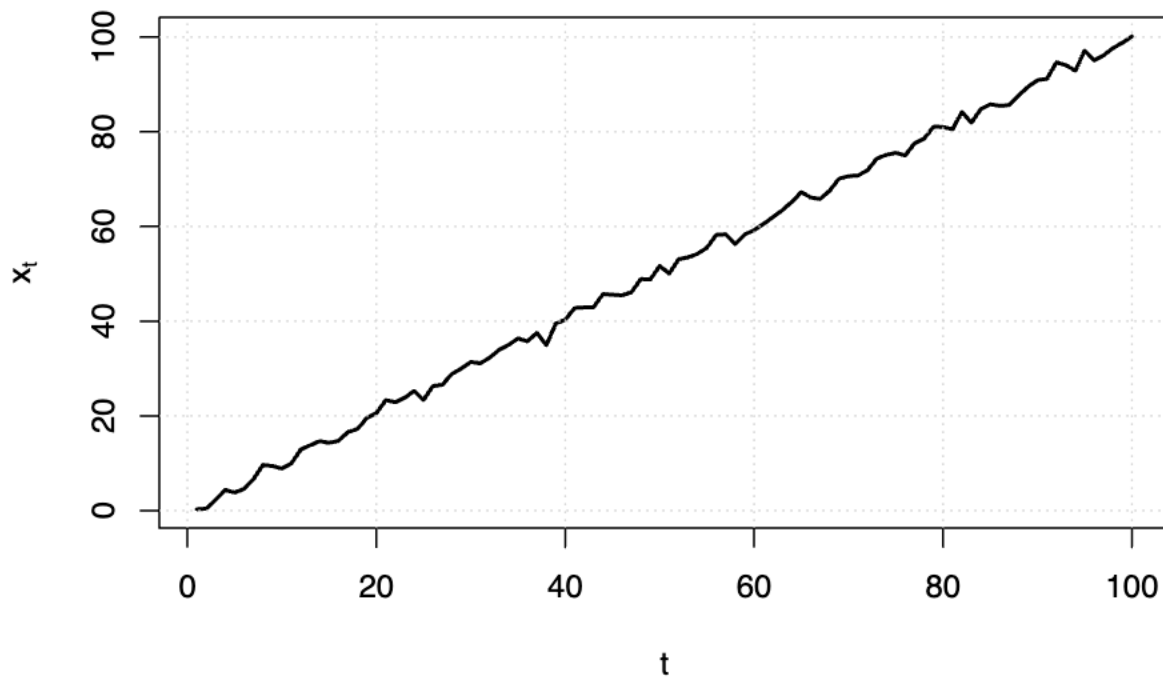
#### --- 0. User-supplied settings -----
student_id <- 1006664584
n          <- 100
beta0      <- 0
beta1      <- 1
sigma_w    <- 1

#### =====
#### === Part (c) : simulate x_t and plot both series =====
#### =====

set.seed(student_id)
t_index <- 1:n
w_t     <- rnorm(n, mean = 0, sd = sigma_w)
x_t     <- beta0 + beta1 * t_index + w_t  #
dx_t    <- diff(x_t)

# ----- plot for part (c) -----
plot(
  t_index, x_t, type = "l", lwd = 2,
  main = expression(paste("Part (c): Simulated series ", x[t], " = ", t, " + ", w[t])),
  xlab = "t", ylab = expression(x[t])
)
grid()
```

Part (c): Simulated series $x_t = t + w_t$



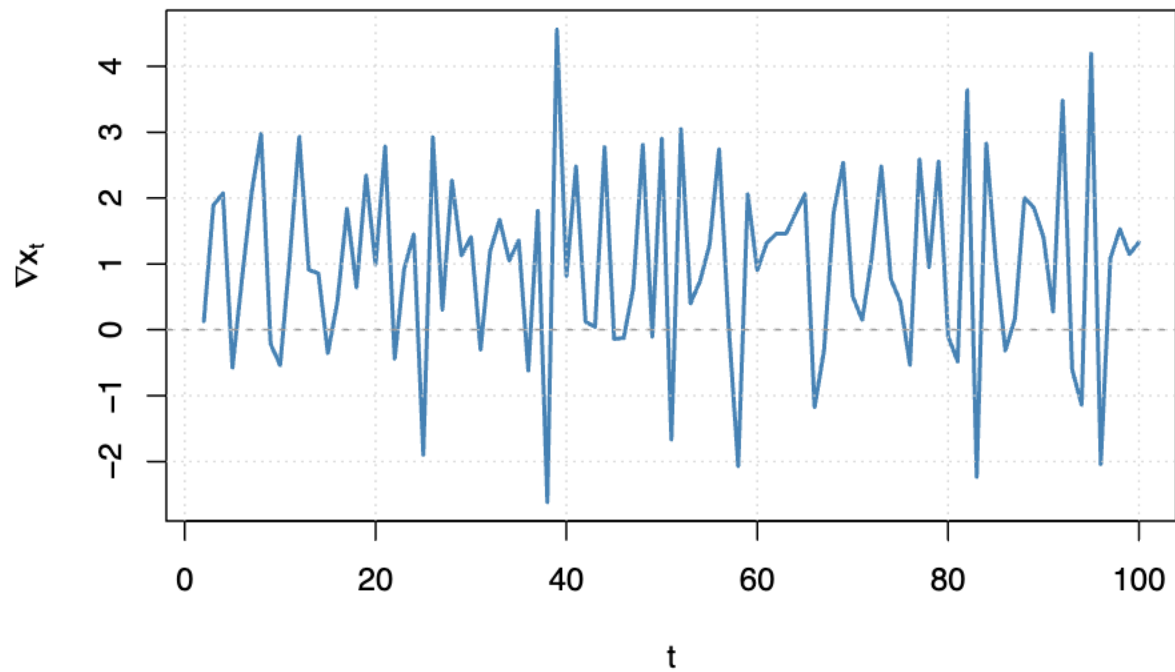
Part (c): The series follows a clear upward linear trend, with white-noise, fluctuations around the line so that it is not stationary.

```
#### =====
#### == Part (d) : first differences and comparison plot ==
#### =====

dx_t <- diff(x_t)

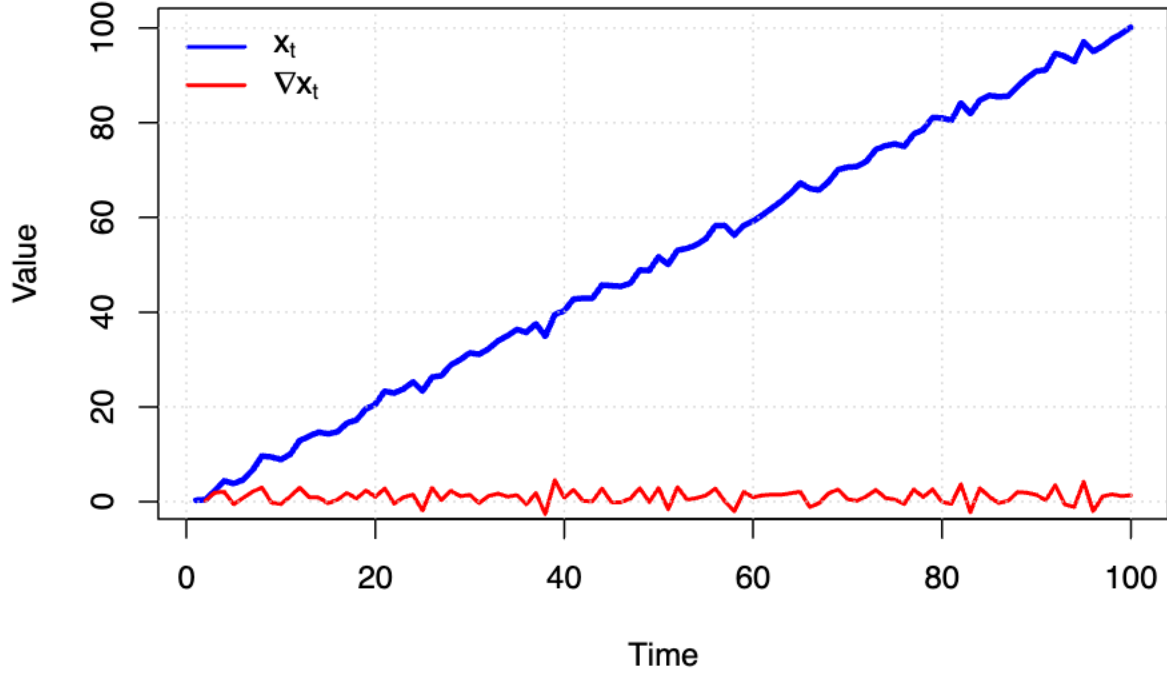
# ----- plot for part (d) -----
plot(
  t_index[-1], dx_t, type = "l", lwd = 2, col = "steelblue",
  main = expression(paste("Part (d): First differences ", nabla, x[t])),
  xlab = "t", ylab = expression(nabla * x[t])
)
abline(h = 0, lty = 2, col = "grey60")
grid()
```

Part (d): First differences ∇x_t



```
# ----- combined plot -----
plot(
  t_index, x_t, type = "l", lwd = 3, col = "blue",
  main = "Original vs. First Difference",
  xlab = "Time", ylab = "Value"
)
lines(t_index[-1], dx_t, lwd = 2, col = "red")
legend(
  "topleft",
  legend = c(expression(x[t]), expression(nabla * x[t])),
  col = c("blue", "red"),
  lwd = 2,
  bty = "n"
)
grid()
```

Original vs. First Difference



Part (d): After differencing, the series oscillates around zero with no discernible trend, which is consistent with stationarity. The first-difference plot reveals that differencing has successfully stripped away the strong linear trend visible in the original x_t series, leaving a sequence of irregular, mean-reverting fluctuations that hover around a constant level of roughly $\beta_1 = 1$. Whereas x_t climbs steadily from near 0 to about 100 with tightly correlated successive points, the differenced series shows no systematic drift, a roughly constant spread throughout the sample, and only weak short-range dependence (a modest negative correlation at lag 1 and essentially zero thereafter). Visually, this translates into the smooth, upward-sloping black line of x_t being replaced by a jagged blue line that oscillates above and below zero with similar amplitude across time. The transformation therefore stabilizes both the mean and the variance, converting a clearly non-stationary process into one whose mean and autocovariance depend only on lag—making it suitable for standard stationary time-series modeling such as ARMA or SARIMA.

(e) **Differencing when w_t is replaced by a *general* stationary process y_t** Let $\{y_t\}$ be stationary with

$$\mathbb{E}(y_t) = \mu_y, \quad \gamma_y(h) = \text{Cov}(y_t, y_{t+h}).$$

Define

$$x_t = \beta_0 + \beta_1 t + y_t, \quad \nabla x_t = x_t - x_{t-1} = \beta_1 + y_t - y_{t-1}.$$

Mean

$$\mathbb{E}(\nabla x_t) = \beta_1 + \mu_y - \mu_y = \beta_1 \quad (\text{constant}).$$

Autocovariance

$$\begin{aligned} \gamma_{\nabla x}(h) &= \text{Cov}(y_t - y_{t-1}, y_{t+h} - y_{t+h-1}) \\ &= \gamma_y(h) - \gamma_y(h-1) - \gamma_y(h+1) + \gamma_y(h) \\ &= 2\gamma_y(h) - \gamma_y(h-1) - \gamma_y(h+1). \end{aligned}$$

Since $\gamma_{\nabla x}(h)$ depends only on the lag h , the differenced series $\{\nabla x_t\}$ is again stationary, even though the original $\{x_t\}$ is not.

A deterministic linear trend renders $\{x_t\}$ non-stationary, but a single first difference removes the trend and yields a stationary series whose second-order properties are easily expressed in terms of those of the innovation process.

```
#####
## Question 2 - Glacial Varve data ##
## (a) heteroskedasticity test ##
## (b) log transform ##
## (c) ACF of log series ##
## (d) first differencing ##
## (e) histogram & ACF of differences ##
#####

# --- 0. Load data -----
if (!requireNamespace("astsa", quietly = TRUE)) {
  install.packages("astsa")
}
library(astsa)          # contains the 'varve' dataset
data(varve)             # load as time-series object
x <- as.numeric(varve)  # raw varve thicknesses

n <- length(x)
mid <- floor(n/2)

# -----
# (a) Compare variance in first and second halves
# -----
var_first <- var(x[1:mid])
var_second <- var(x[(mid + 1):n])

cat(sprintf(
  "(a) Sample variance 1-%d: %8.3f\n", mid, var_first))

## (a) Sample variance 1-317: 133.457

cat(sprintf(
  " Sample variance %d-%d: %8.3f\n\n", mid + 1, n, var_second))

## Sample variance 318-634: 594.490
```

The sample calculation shows that the dispersion of glacial-varve thicknesses changes dramatically over time: the first 317 observations have a sample variance of about 133.457, while the second 317 observations have a variance of 594.490—more than four times larger. Because variance is a measure of average squared deviation from the mean, such a pronounced jump implies that the spread (and hence the uncertainty) of the series is not constant across the record. In other words, the process is heteroskedastic: earlier layers display relatively modest fluctuations around their mean, whereas later layers exhibit much larger swings, indicating that the variability of sediment deposition increased substantially in the second half of the sequence.

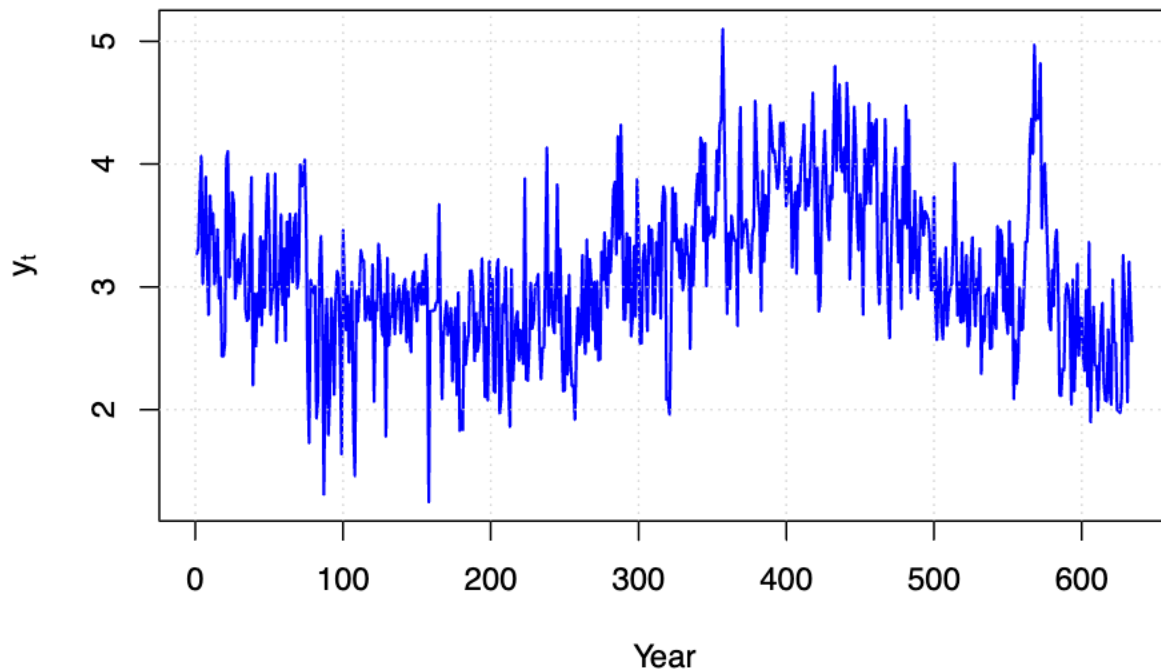
```

# -----
# (b) Log-transform, plot, eyeball variance stabilization
# -----
# Handle zeros: drop them (usual in this dataset)
y <- log(x)
y[!is.finite(y)] <- NA
yt <- ts(na.omit(y))

plot(yt, type = "l", col = "blue", lwd = 1.5,
     main = "(b) Log-transformed varve series  y_t = log x_t",
     ylab = expression(y[t]), xlab = "Year")
grid()

```

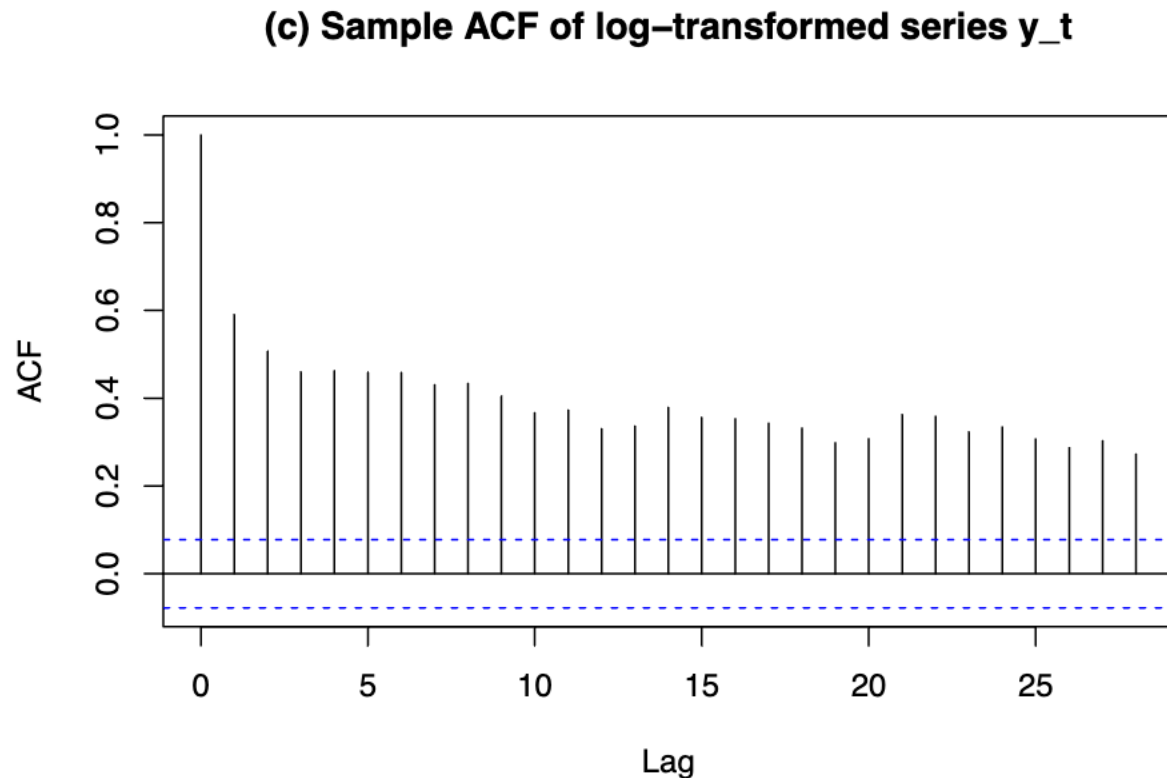
(b) Log-transformed varve series $y_t = \log x_t$



Variance of y_t (first half): 0.271 \ Variance of y_t (second half): 0.451

By comparing the pre- and post-transformation variances it is clear that the log transform has greatly stabilised the series' variance. The log transformation compresses the large varve values in the second half of the record much more than the smaller values in the first half, so the amplitude of the oscillations now looks roughly comparable throughout the 634 observations: peaks and troughs fluctuate within a band of 2–5 log-units instead of expanding from tens to hundreds of original units. A quick numerical check confirms the visual impression: the sample variance falls from about 133.457 to 0.271 in the second half and from 594.490 to 0.451 in the first half, so the ratio of second-half to first-half variance drops from 4.45 to 1.66. Although the transformed series still shows slow undulations in level (e.g., a bump around years 350–500), the dramatic heteroskedasticity evident in the raw data has been substantially reduced; the remaining variation is much more homoscedastic and suitable for further modelling steps such as ACF analysis and differencing.

```
# -----
# (c) Sample ACF of the log-transformed series
# -----
acf(yt, main = "(c) Sample ACF of log-transformed series y_t")
```

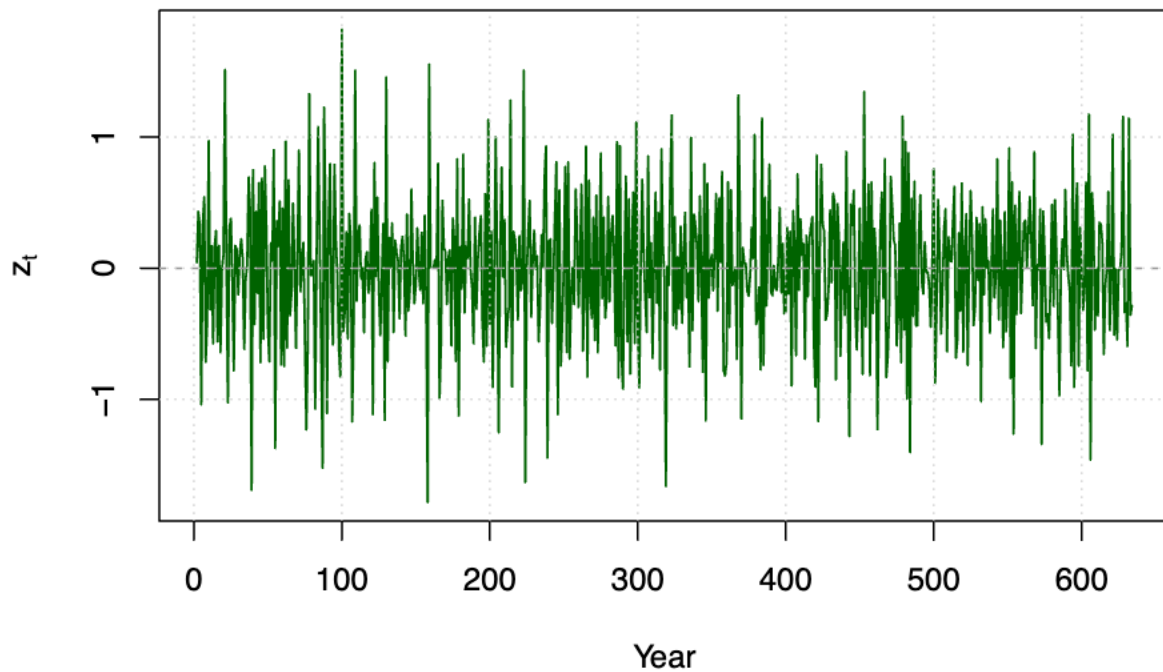


The sample ACF of the log-transformed varve thicknesses begins near 1.0 at lag 1 and then declines very slowly, remaining positive and well outside the blue 95 % bounds for at least the first 25–30 lags. This pattern of a high first-lag correlation followed by a gradual, almost geometric decay is classic evidence of strong persistence: today’s value is highly predictive of tomorrow’s, and shocks dissipate only slowly over time. Because the bars never cross into negative territory and the decay is so gentle, the series still looks “long-memory” or nearly non-stationary even after the variance-stabilising log transform—behaviour consistent with an AR(1) process whose coefficient is close to 1 or with an integrated (ARIMA) process that has not yet been differenced. In practical terms, the ACF suggests that further transformation—most commonly a first difference—will be required to remove the remaining serial dependence before fitting short-memory ARMA models or conducting inference that assumes weak dependence.

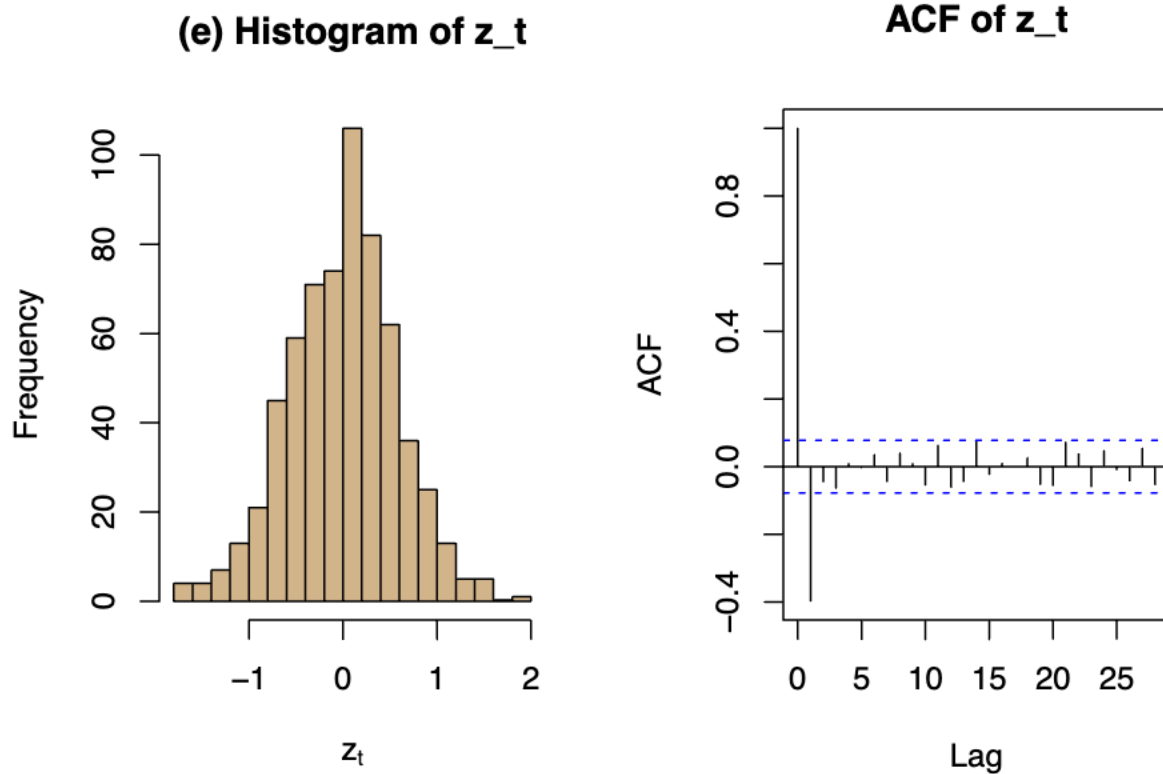
```
# -----
# (d) First differences of the log series
# -----
zt <- diff(yt)

plot(zt, type = "l", col = "darkgreen", lwd = 1.2,
     main = "(d) First-differenced log series z_t",
     ylab = expression(z[t]), xlab = "Year")
abline(h = 0, lty = 2, col = "gray60")
grid()
```


(d) First-differenced log series z_t



```
# -----  
# (e) Histogram and ACF of  $z_t$   
# -----  
par(mfrow = c(1, 2))           # side-by-side plots  
hist(zt, breaks = "FD", col = "tan",  
     main = "(e) Histogram of  $z_t$ ", xlab = expression(z[t]))  
acf(zt, main = "ACF of  $z_t$ ")
```



```
par(mfrow = c(1, 1))          # reset

# Quick numerical check of variance stability after log
mid2 <- floor(length(yt)/2)
cat(sprintf("\nVariance of y_t (first half) : %8.3f\n", var(yt[1:mid2])))

##
## Variance of y_t (first half) :    0.271

cat(sprintf("Variance of y_t (second half): %8.3f\n", var(yt[(mid2+1):length(yt)])))

## Variance of y_t (second half):    0.451
```

The differenced log-varve series $z_t = y_t - y_{t-1}$ behaves like a weakly stationary process. The histogram of z_t —the first differences of the log-varve series—clusters tightly around zero and is roughly bell-shaped, with only mild skewness and no evidence of heavy tails; this suggests the series has a constant mean and a stable variance across time. The accompanying sample ACF corroborates that impression: apart from the inevitable unit spike at lag 0, the only notable bar is a small negative coefficient at lag 1 (a common artefact of differencing), while all subsequent lags lie well inside the $\pm 2/\sqrt{N}$ confidence bounds and fluctuate randomly around zero. Such a “flat” ACF indicates the absence of long-range dependence and implies that the second-order properties (mean, variance, autocovariance) do not change with time. Taken together, the nearly normal histogram and the rapidly dying ACF provide strong empirical evidence that the differenced log series z_t is weakly stationary and behaves much like white noise, making it an appropriate starting point for modelling with short-memory processes such as an ARMA model (often an ARMA(0,1) or even pure noise).

3. (a) Autocovariance function $\gamma(h)$ of an MA(1) model

Assume the differenced log-varve series can be written as

$$z_t = \mu + w_t + \theta w_{t-1}, \quad w_t \stackrel{\text{i.i.d.}}{\sim} (0, \sigma_w^2),$$

with mutually independent innovations w_t .

$$\gamma(h) = \text{Cov}(z_t, z_{t+h}) = \mathbb{E}[(z_t - \mu)(z_{t+h} - \mu)], \quad h \in \mathbb{Z}.$$

$h = 0$:

$$\gamma(0) = \text{Var}(z_t) = \text{Var}(w_t + \theta w_{t-1}) = \sigma_w^2 + \theta^2 \sigma_w^2 = (1 + \theta^2) \sigma_w^2.$$

$|h| = 1$:

$$\gamma(1) = \text{Cov}(z_t, z_{t+1}) = \text{Cov}(w_t + \theta w_{t-1}, w_{t+1} + \theta w_t) = \theta \sigma_w^2,$$

and by symmetry $\gamma(-1) = \gamma(1)$.

$|h| > 1$: If $|h| > 1$, the terms w_t and w_{t+h} share no common innovation, hence

$$\gamma(h) = 0, \quad |h| > 1.$$

$$\gamma(h) = \begin{cases} (1 + \theta^2) \sigma_w^2, & h = 0, \\ \theta \sigma_w^2, & |h| = 1, \\ 0, & |h| > 1. \end{cases}$$

3. (b) Autocorrelation function $\rho(h)$

Because $\rho(h) = \gamma(h)/\gamma(0)$ with $\gamma(0) = (1 + \theta^2) \sigma_w^2$,

$$\rho(h) = \begin{cases} 1, & h = 0, \\ \frac{\theta}{1 + \theta^2}, & |h| = 1, \\ 0, & |h| > 1. \end{cases}$$

Thus an MA(1) process has non-zero autocovariance (and correlation) only at lags 0 and ± 1 ; all higher-order lags are uncorrelated.

```
#####
## (c)
#####

gamma0_hat <- var(zt)

acf_zt <- stats::acf(zt, lag.max = 1, plot = FALSE)
rho1_hat <- as.numeric(acf_zt$acf[2])

# ---- LaTeX-friendly output -----
```

```
# Use results = "asis" in the chunk header if you want the
# lines to be passed straight through to the .tex file.
```

```
cat(sprintf("\nEstimated $\hat{\gamma}(0)$ = %.3f\\n", gamma0_hat))
```

```
##
## Estimated $\hat{\gamma}(0)$ = 0.332\\
```

```
cat(sprintf("Estimated $\hat{\rho}(1)$ = %.3f\\n", rho1_hat))
```

```
## Estimated $\hat{\rho}(1)$ = -0.397
```

$$\hat{\gamma}(0) = 0.332, \quad \hat{\rho}(1) = -0.397.$$

3. (d) Method-of-moments estimates of θ and σ_w^2

From parts (a)–(b) the theoretical moments of an MA(1) process are

$$\gamma(0) = (1 + \theta^2) \sigma_w^2, \quad \rho(1) = \frac{\theta}{1 + \theta^2}.$$

Equating these to the empirical values obtained in Question 2(c),

$$\hat{\gamma}(0) = 0.332, \quad \hat{\rho}(1) = -0.397,$$

gives the moment equations

$$\begin{cases} 0.332 = (1 + \theta^2) \sigma_w^2, \\ -0.397 = \frac{\theta}{1 + \theta^2}. \end{cases}$$

Solving for θ

From the second equation, multiply both sides by $1 + \theta^2$:

$$-0.397(1 + \theta^2) = \theta \implies -0.397\theta^2 - \theta - 0.397 = 0.$$

This quadratic can be written as

$$(\hat{\rho}) \theta^2 - \theta + \hat{\rho} = 0, \quad \text{with } \hat{\rho} = -0.397.$$

Using the quadratic formula,

$$\theta = \frac{1 \pm \sqrt{1 - 4\hat{\rho}^2}}{2\hat{\rho}} = \frac{1 \pm \sqrt{1 - 4(0.397)^2}}{2(-0.397)} = \frac{1 \pm 0.608}{-0.794}.$$

Hence θ has two roots:

$$\theta_1 \approx -2.03, \quad \theta_2 \approx -0.494.$$

For an MA(1) model to be **invertible** we must have $|\theta| < 1$; therefore we choose

$$\hat{\theta} = -0.494.$$

Solving for σ_w^2

Substitute $\hat{\theta}$ into the first moment equation:

$$\hat{\sigma}_w^2 = \frac{\hat{\gamma}(0)}{1 + \hat{\theta}^2} = \frac{0.332}{1 + (-0.494)^2} = \frac{0.332}{1 + 0.244} \approx 0.267.$$

$$\hat{\sigma}_w^2 \approx 0.267$$

Thus, by the method of moments the estimated MA(1) parameters for the differenced log-varve series are

$$\hat{\theta} \approx -0.494, \quad \hat{\sigma}_w^2 \approx 0.267.$$