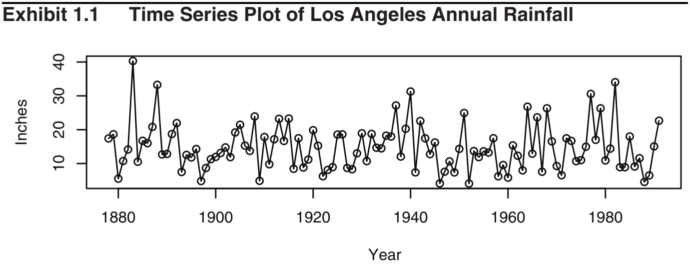
Ch 2: Fundamental concepts

- What are time series and stochastic processes?
- How to describe stochastic processes?
  - Mean and autocovariance (autocorrelation) functions.
  - Stationarity.
- Material: Ch 2, excluding the example on the random cosine wave (p.18+).

Time series

- A time series is a set of observations, each one being recorded at a specific time t.



- We focus on discrete time series, where observation times are indexed by $t = 0, \pm 1, \pm 2, \pm 3, \ldots$.
- We will consider each observed value y_t to be a realization of the random variable Y_t (e.g., a draw from a standard normal distribution).
  - Goal: describe the distribution of Y_t!
- Note: We will often stick to the notation of the book, whereby Y_t is used to denote the random variable as well as the observation.

Time series and stochastic processes

- The sequence of random variables $\{Y_t : t = 0, \pm 1, \pm 2, \pm 3, \ldots\}$ is called a stochastic process, also referred to as a time series process.
- Examples of (simple) stochastic processes in Ch 2:
  White noise, random walk, moving average.
- Our goal in time series analysis:
  Given an observed time series, find out what time series process(es) could have resulted in the observed time series; use a stochastic process as a model for the time series of interest.
- Step 1: learn about stochastic processes and their properties.

Stochastic process, example 0: White noise

- White noise = time series of random variables $\{e_t : t = 0, \pm 1, \pm 2, \pm 3, \ldots\}$ with
  - identical distributions for each e_t, with $E(e_t) = 0$ and $Var(e_t) = \sigma_e^2$ (constant),
  - zero correlation between the e_t's, $cor(e_t, e_p) = 0$ for time $t \neq p$.
- In book 2 (Brockwell and Davis), white noise is defined as above with uncorrelated random variables. We will use the more strict definition as used in our book, which is, that the e_t's are independent.
- Throughout the book/course, $\{e_t\}$ always refers to white noise.
- Short notation: $e_t \sim WN(0, \sigma_e^2)$
- White noise is used as a building block for many more complicated stochastic processes.

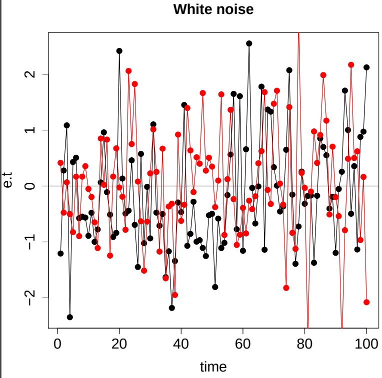
Simulating white noise in R

- We can simulate and plot an observed series of white noise, $e_t \sim WN(0, \sigma_e^2)$ for $t = 1, \ldots, 100$ in R.
  - We need a probability distribution for simulating the e_t's, let's use a normal distribution: $e_t \sim N(0, \sigma_e^2)$.

Rcode

```
sigma.e <- 1
n <- 100
time <- seq(1, n)

set.seed(1234)
e.t <- rnorm(n, 0, sigma.e)
```



Stochastic process, example 1: Random walk

- Random walk = time series $\{Y_t : t = 1, 2, 3, \ldots\}$:

$$\begin{aligned} Y_1 &= e_1, \\ Y_2 &= e_1 + e_2, \\ &\vdots \\ Y_t &= e_1 + e_2 + \ldots + e_t \end{aligned}$$

  where $e_1, e_2, \ldots$ is white noise.
- Equivalently

$$\begin{aligned} Y_1 &= e_1, \\ Y_t &= Y_{t-1} + e_t \text{ for } t > 1, \\ e_t &\sim WN(0, \sigma_e^2), \end{aligned}$$

  with $Y_1 = e_1$.
- Random walk process used as a model for stock prices, projecting changes in fertility, ...

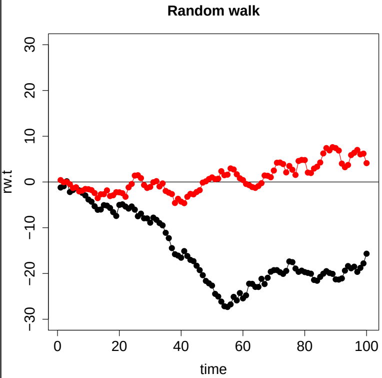
Simulating a random walk

- Based on a simulation of white noise (e.g., $e_t \sim N(0, \sigma_e^2)$ as before), we can calculate and plot the corresponding random walk $\{Y_t : t = 1, 2, 3, \ldots\}$ with

$$Y_t = Y_{t-1} + e_t \text{ for } t > 1 \text{ and } Y_1 = e_1.$$

R-code (continued)

```
set.seed(1234)
e.t <- rnorm(n, 0, sigma.e)
rw.t <- cumsum(e.t)
plot(rw.t ~ time, type = "l",
main = "Random walk",
ylim  = c(-30,30))
```



Stochastic process, example 2: A moving average

- Moving average = time series $\{Y_t : t = 0, \pm 1, \pm 2, \pm 3, \ldots\}$:

$$Y_t = \frac{e_t + e_{t-1}}{2},$$

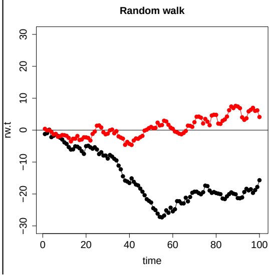
  where $e_t \sim WN(0, \sigma_e^2)$.
- Construct your own simulations in R!

Describing stochastic processes

- How to describe stochastic or time series processes?
  - The multivariate distribution of the Y_t's fully specifies a time series process.
  - (Easier) starting point: describe means and (co)variances (1st and 2nd order moments).
  - Note: see Appendix A for a general review of mean and (co)variance.
- The mean function for a time series process is defined as:

$$\mu_t = E(Y_t) \text{ for } t = 0, \pm 1, \pm 2, \pm 3, \ldots. \tag{1}$$

What is the mean function for random walk with $Y_t = Y_{t-1} + e_t$ and $Y_1 = e_1$?



Autocovariance function

- The autocovariance function (ACVF) describes the covariance between the Y_t's for any two points in time, e.g, Y_1 and Y_2.
- Autocovariance function (ACVF) $\gamma_{t,s}$:

$$\begin{aligned}\gamma_{t,s} &= Cov(Y_t, Y_s), \text{ for } s, t = 0, \pm 1, \pm 2, \pm 3, \ldots \\ &= E[(Y_t - \mu_t)(Y_s - \mu_s)], \\ &= E(Y_t Y_s) - \mu_t \mu_s.\end{aligned}$$

- Q: What is/how do we call $\gamma_{t,t}$?

$$\gamma_{t,t} = Cov(Y_t, Y_t) = Var(Y_t)$$

- Q: Is $\gamma_{t,s} = \gamma_{s,t}$?

$$\gamma_{t,s} = Cov(Y_t, Y_s) = Cov(Y_s, Y_t) = \gamma_{s,t}.$$

- Q: What is $\gamma_{s,t}$ for the random walk?

Autocovariance function for the random walk

- Variance for $t \geq 1$:

$$\begin{aligned} Var(Y_t) &= Var(e_1 + e_2 + \ldots + e_t), \\ &= Var(e_1) + Var(e_2) + \ldots + Var(e_t), \text{ (why?)} \\ &= t \cdot \sigma_e^2. \end{aligned}$$

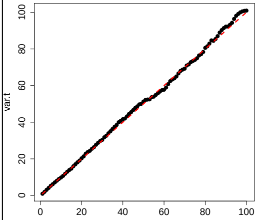
- Interpretation?
- Side question: Can we quickly do a check whether this is correct in R?

Checking the expression for the variance

- How to check whether the expression $\gamma_{t,t} = var(Y_t) = t\sigma_e^2$ for the random walk is correct, through simulation?
  - Simulate many trajectories, e.g. $Y_1^{(k)}, Y_2^{(k)}, \ldots, Y_n^{(k)}$ for $k = 1, 2, \ldots, 1000$.
  - For each t, calculate the sample variance of $(Y_t^{(1)}, Y_t^{(2)}, \ldots, Y_t^{(1000)})$ and compare to $t\sigma_e^2$: if the expression $t\sigma_e^2$ is correct, we do not expect to see differences between the sample variance and $t\sigma_e^2$.

```
K <- 1000 # no of trajectories
rw.tk <- matrix(NA, n, K)
set.seed(1234)
for (k in 1:1000){
  e.t <- rnorm(n, 0, sigma.e)
  rw.tk[,k] <- cumsum(e.t)
}
var.t <- apply(rw.tk,1,var)
plot(var.t ~ time)
```

```
lines(time*sigma.e^2 ~ time,
 col = 2)
> round(var(rw.tk[20,]))
[1] 20
```



Autocovariance continued

- Autocovariance function (ACVF) $\gamma_{t,s}$:

$$\gamma_{t,s} = Cov(Y_t, Y_s), \text{ for } s, t = 0, \pm1, \pm2, \pm3, \ldots$$

- Autocovariance for the random walk $1 \leq t \leq s$:

$$\begin{aligned}
\gamma_{t,s} &= Cov(Y_t, Y_s), \\
&= Cov(e_1 + e_2 + \ldots + e_t, e_1 + e_2 + \ldots + e_t + e_{t+1} + \ldots + e_s), \\
&= \sum_{j=1}^{t} \sum_{i=1}^{s} Cov(e_j, e_i) = \sum_{j=1}^{t} Cov(e_j, e_j) + \sum_{j=1}^{t} \sum_{i \neq j} Cov(e_j, e_i), \\
&= t\sigma_e^2 + 0.
\end{aligned}$$

- Interpretation? (A bit) easier for the autocorrelation.

Autocorrelation function

- The autocorrelation function (ACF) $\rho_{t,s}$ gives the autocorrelation between Y_t for any two times t and s:

$$\begin{aligned}\rho_{t,s} &= Corr(Y_t, Y_s) \text{ for } s, t = 0, \pm 1, \pm 2, \pm 3, \ldots \\ &= \frac{Cov(Y_t, Y_s)}{\sqrt{Var(Y_t)Var(Y_s)}}, \\ &= \frac{\gamma_{t,s}}{\sqrt{\gamma_{t,t}\gamma_{s,s}}}.\end{aligned}$$

- Q: What is $\rho_{t,t}$?
- For random walk for $1 \leq t \leq s$:

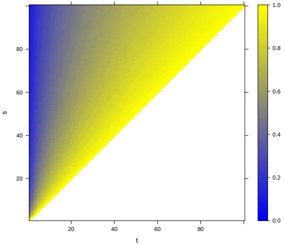
$$\rho_{t,s} = \frac{\gamma_{t,s}}{\sqrt{\gamma_{t,t}\gamma_{s,s}}} = \frac{t\sigma_e^2}{\sqrt{t \cdot s}\sigma_e^2} = \sqrt{t/s}.$$

- Interpretation? Does correlation in/decrease with s and/or t?
- Plot the true autocorrelation function in R!

Autocorrelation function for the random walk

- For random walk for $1 \leq t \leq s$: $\rho_{t,s} = \sqrt{t/s}$.
- Interpretation? Does correlation in/decrease with s and/or t?

R-code (main ideas)




```
cor.ts <- matrix(NA, n, n)
for (t in 1:n){
  for (s in t:n){
    cor.ts[t,s] <- sqrt(t/s)
  }
}
levelplot(cor.ts, xlab = "t", ylab = "s", ylab = "Time")
```

Stochastic process, example 2: A moving average

- For the moving average $\{Y_t : t = 0, \pm 1, \pm 2, \pm 3, \dots\}$:

$$Y_t = \frac{e_t + e_{t-1}}{2},$$

where $e_t \sim WN(0, \sigma_e^2)$, verify for yourself (see answer in book):

$$\mu_t = 0,$$
$$\gamma_{t,s} = \begin{cases} 0.5\sigma_e^2 & \text{for } |t-s| = 0, \\ 0.25\sigma_e^2 & \text{for } |t-s| = 1, \\ 0 & \text{otherwise.} \end{cases}$$

- Equivalently with $s = t + k$:

$$\gamma_{t,t+k} = \begin{cases} 0.5\sigma_e^2 & \text{for } k = 0, \\ 0.25\sigma_e^2 & \text{for } |k| = 1, \\ 0 & \text{otherwise,} \end{cases}$$

and $\rho_{t,t+k} = 1$ for $k = 0$, $\rho_{t,t+k} = 0.5$ for $k = \pm 1$ and 0 otherwise (much easier to draw!).

- Do μ_t or $\gamma_{t,t+k}$ depend on t?

Stationarity

- A stochastic process $\{Y_t\}$ is said to be **weakly** (or **second-order**) stationary if
  1. μ_t is constant over time,
  2. $\gamma_{t,t-k} = \gamma_{0,k}$, for all time t and lag k.
- Note: In the book/course (generally), the term *stationary* without additional info refers to this specific form of weak stationary.
- Is the moving average of e_t's stationary?
- Is the random walk stationary?

$$
\begin{aligned}
\mu_t &= 0, \\
\gamma_{t,s} &= t\sigma_e^2 \text{ for } 1 \le t \le s.
\end{aligned}
$$

- Is the differenced random walk, $\nabla Y_t = Y_t - Y_{t-1} = e_t$ stationary? (to be continued!).

Strict stationarity

- A process $\{Y_t\}$ is said to be **strictly stationary** if the joint distribution of $Y_1, Y_2, \dots, Y_{t_n}$ is the same as the joint distribution of $Y_{1-k}, Y_{2-k}, \dots, Y_{t_n-k}$ for all choices of time points $t_1, t_2, \dots, t_n$ and all choices of time lag k.
- Is white noise strictly stationary?

$$
\begin{aligned}
& Pr(e_{t_1} \le x_1, e_{t_2} \le x_2, \dots, e_{t_n} \le x_n) \\
&= Pr(e_{t_1} \le x_1) Pr(e_{t_2} \le x_2) \cdots Pr(e_{t_n} \le x_n) \qquad \text{(by independence)} \\
&= Pr(e_{t_1-k} \le x_1) Pr(e_{t_2-k} \le x_2) \cdots Pr(e_{t_n-k} \le x_n) \\
& \qquad \qquad \text{(identical distributions)} \\
&= Pr(e_{t_1-k} \le x_1, e_{t_2-k} \le x_2, \dots, e_{t_n-k} \le x_n) \qquad \text{(by independence)}
\end{aligned}
$$

Yes!

Summary of Ch 2: Fundamental concepts

- We discussed definitions of time series and stochastic processes, as well as examples of stochastic processes (time series models).
- We learned how to describe stochastic processes with their mean and autocovariance (autocorrelation) functions, and when a process is said to be stationarity.
- Next: Ch 4 (we skip Ch 3):
  - More time series models: moving average and autoregressive processes.