Covariance & Correlation

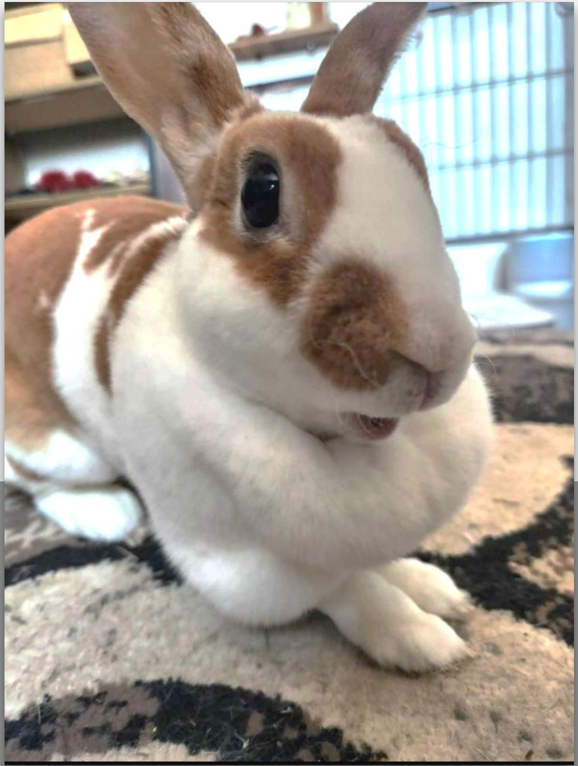
A Time Series Analyst’s Crash Course

STAT 464 / 864 | Fall 2024

Discrete Time Series Analysis

Skyepaphora Griffith, Queen’s University

We learned something on Wednesday, in Time Series 🕛



The S1 method

Eliminates seasonality

Gives us a model for the data's noise, Y_t

What did we tell quin??

Time to get Random

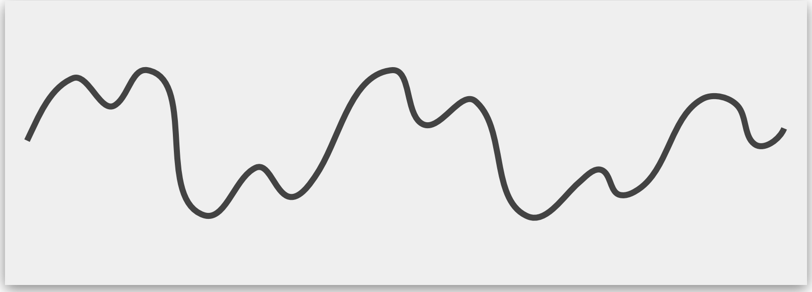
$$X_t = \boxed{m_t} + \boxed{s_t} + \boxed{Y_t} \qquad (\star)$$

- 🕑 We’ve **removed the trend and seasonal** components.
- 🕑 Still gotta model the **noise!**

In time series, we’re interested in **temporal** or **serial dependence**

If there is none? Then it’s **noise.** ¯_(ツ)_/¯

We often measure this dependence using **correlation** (std. **covariance**)



Space for notes

Covariance | Definition

For random variables **X** and **Y**:

$$\mathrm{Cov}(X, Y) = \mathrm{E}\Big[(X - \mathrm{E}[X])(Y - \mathrm{E}[Y])\Big]$$

$\mathrm{Cov}(X, Y) > 0$ **Positively** correlated

$\mathrm{Cov}(X, Y) < 0$ **Negatively** correlated

$\mathrm{Cov}(X, Y) = 0$ **Not** correlated (**un**correlated)

Space for notes

Side Note: Expectation is Linear!

$$\mathrm{E}\left[\sum_{i=1}^{N} c_i X_i\right] = \sum_{i=1}^{N} \mathrm{E}[c_i X_i] = \sum_{i=1}^{N} c_i \mathrm{E}[X_i]$$

Covariance | Alternative Formula

$$
\begin{aligned}
\operatorname{Cov}(X,Y) &= \mathrm{E}\big[(X-\mathrm{E}[X])(Y-\mathrm{E}[Y])\big] \\
&= \mathrm{E}\big[XY - \big(X\cdot\mathrm{E}[Y]\big) - \big(Y\cdot\mathrm{E}[X]\big) + \mathrm{E}[X]\,\mathrm{E}[Y]\big] \\
&= \mathrm{E}[XY] - \mathrm{E}\big[X\cdot\mathrm{E}[Y]\big] - \mathrm{E}\big[Y\cdot\mathrm{E}[X]\big] + \mathrm{E}[X]\,\mathrm{E}[Y] \\
&= \mathrm{E}[XY] - \mathrm{E}[Y]\,\mathrm{E}[X] - \mathrm{E}[X]\,\mathrm{E}[Y] + \mathrm{E}[X]\,\mathrm{E}[Y] \\
&= \mathrm{E}[XY] - \mathrm{E}[X]\mathrm{E}[Y]
\end{aligned}
$$

Covariance | Alternative Formula

$$\mathrm{Cov}(X, Y) = \mathrm{E}[XY] - \mathrm{E}[X]\mathrm{E}[Y]$$

Covariance Properties | Overview

C1) Symmetry $\text{Cov}(X, Y) = \text{Cov}(Y, X)$

C2) RVs are uncorrelated with constants $\text{Cov}(c, X) = 0$

C3) Covariance is linear in both arguments

C4) Variance = Covariance of RV with itself

C5) Independent $\overset{\Longrightarrow}{\not\Longleftarrow}$ Uncorrelated

C6) Cauchy-Schwarz: $\text{Cov}(X, Y)^2 \leq \text{Var}(X)\text{Var}(Y)$

Space for notes 🕛 🕑 🕞

C3) Linearity in both arguments

$$\mathrm{Cov}\left(\sum_{i=1}^{m} a_i X_i, \sum_{j=1}^{n} b_j Y_j\right) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_i b_j \mathrm{Cov}(X_i, Y_j)$$

$$\begin{aligned}\mathrm{Cov}(a_1X_1 + a_2X_2, b_1Y_1 + b_2Y_2) &= a_1b_1\mathrm{Cov}(X_1, Y_1)\\ &+ a_1b_2\mathrm{Cov}(X_1, Y_2)\\ &+ a_2b_1\mathrm{Cov}(X_2, Y_1)\\ &+ a_2b_2\mathrm{Cov}(X_2, Y_2)\end{aligned}$$

C4) Variance is Covariance

$$\mathrm{Cov}(X, X) = \mathrm{E}[X^2] - \mathrm{E}[X]^2$$

$$= \mathrm{Var}[X]$$

C4) Variance... like... really *wants* to be linear, but no

$$\mathrm{Var}\left(\sum_{i=1}^{n} a_i X_i\right) = \mathrm{Cov}\left(\sum_{i=1}^{n} a_i X_i, \sum_{j=1}^{n} a_j X_j\right)$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j \mathrm{Cov}(X_i, X_j)$$

$$= \sum_{i=1}^{n} a_i^2 \mathrm{Var}(X_i) + 2\sum_{i=1}^{n}\sum_{\substack{j=1 \\ j<i}}^{n} a_i a_j \mathrm{Cov}(X_i, X_j)$$

For uncorrelated RVs: variance of the sum = sum of the variance

C5a) Independent $\implies$ Uncorrelated

$$\mathrm{Cov}(X, Y) = \mathrm{E}[XY] - \mathrm{E}[X]\mathrm{E}[Y]$$

$$= 0$$

C5b) Independent $\not\Leftarrow$ Uncorrelated

Let X have a symmetric distribution about 0

Assume $\mathrm{E}[X^k]$ exists for $k \leq 3$

$\mathrm{E}[X^k] = 0$ for all odd k, since such X^k are odd functions about 0

Let $Y = X^2$

$$\begin{aligned}\mathrm{Cov}(X, Y) &= \mathrm{E}[X \cdot X^2] - \mathrm{E}[X]\mathrm{E}[X^2] \\ &= \boxed{\mathrm{E}[X^3]} - \boxed{\mathrm{E}[X]}\mathrm{E}[X^2] = 0\end{aligned}$$

Space for notes 🕛 🕑 🕒

C6) Cauchy-Schwarz Inequality

$$\boxed{0 \leq} \mathrm{E}[(X - \lambda Y)^2] = \mathrm{E}[X^2 - 2\lambda XY + \lambda^2 Y^2]$$
$$= \boxed{\mathrm{E}[X^2] - 2\lambda \mathrm{E}[XY] + \lambda^2 \mathrm{E}[Y^2]}$$

Minimize: take derivative with respect to λ

$$0 = -2\mathrm{E}[XY] + 2\lambda \mathrm{E}[Y^2]$$

$$\lambda = \frac{E[XY]}{E[Y^2]}$$

Space for notes 🕛 🕑 🕞

C6) Cauchy-Schwarz Inequality

$$\boxed{0 \leq \mathrm{E}[X^2] - 2\lambda\mathrm{E}[XY] + \mathrm{E}[\lambda^2 Y^2]}$$

Plug in λ

$$\lambda = \frac{E[XY]}{E[Y^2]}$$

$$= \mathrm{E}[X^2] - \frac{2\mathrm{E}[XY]}{\mathrm{E}[Y^2]}\mathrm{E}[XY] + \frac{\mathrm{E}[XY]^2}{\mathrm{E}[Y^2]^2}\mathrm{E}[Y^2]$$

$$= \mathrm{E}[X^2] - \frac{2\mathrm{E}[XY]^2}{\mathrm{E}[Y^2]} + \frac{\mathrm{E}[XY]^2}{\mathrm{E}[Y^2]}$$

$$= \boxed{\mathrm{E}[X^2] - \frac{\mathrm{E}[XY]^2}{\mathrm{E}[Y^2]}}$$

Space for notes 🕛 🕑 🕞

C6) Cauchy-Schwarz Inequality

$$\boxed{0 \leq \mathrm{E}[X^2] - \frac{\mathrm{E}[XY]^2}{\mathrm{E}[Y^2]}}$$

$$\mathrm{E}[XY]^2 \leq \mathrm{E}[X^2]\mathrm{E}[Y^2]$$

$$\mathrm{E}\big[(X - \mathrm{E}[X])(Y - \mathrm{E}[Y])\big]^2 \leq \mathrm{E}\big[(X - \mathrm{E}[X])^2\big]\mathrm{E}\big[(Y - \mathrm{E}[Y])^2\big]$$

$$\mathrm{Cov}(X, Y)^2 \leq \mathrm{Var}(X)\mathrm{Var}(Y)$$

Space for notes 🕛 🕑 🕒

Covariance vs. Correlation

Drawback of using covariance as measure of dependence:

It depends on the units of the RVs.

Example:

If **X, Y** measured are in **cm** → **10X** and **10Y** are same quantities in **mm**

But notice: **Cov[10X, 10Y] = 100 Cov[X,Y]** → cov is not standardized.

Don’t want the **measure** of dependence to change based on units

Space for notes 🕛 🕑 🕒

Correlation | Definition & Properties

Correlation coefficient:

$$\rho(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}$$

1) Symmetry (follows from symmetry of cov.)

2) Unitless, real valued on **[-1, 1]**

3) $\rho = \pm 1 \iff$ Y is a linear function of **X,**

   ie) **Y = a + bX**

Space for notes 🕛 🕑 🕞

③ $\rho(X,Y)=\pm 1$ if and only if Y is a linear function of X.

Proof Suppose $Y=a+bX$ for some nonzero constant b.

Then $\text{Cov}(X,Y)=\text{Cov}(X,a+bX)$
$=\text{Cov}(X,a)+b\,\text{Cov}(X,X)$
$=b\,\text{Var}(X)$

$\text{Var}(Y)=\text{Var}(a+bX)=b^2\text{Var}(X)$

So $\rho(X,Y)=\dfrac{b\,\text{Var}(X)}{\sqrt{\text{Var}(X)\,b^2\,\text{Var}(X)}}=\dfrac{b}{|b|}=\pm 1$

Conversely, suppose $\rho(X,Y)=+1$. Let $\sigma_X^2=\text{Var}(X)$ and $\sigma_Y^2=\text{Var}(Y)$. Then

$$\text{Var}\left(\frac{X}{\sigma_X}-\frac{Y}{\sigma_Y}\right)=\frac{\text{Var}(X)}{\sigma_X^2}+\frac{\text{Var}(Y)}{\sigma_Y^2}-2\,\text{Cov}\left(\frac{X}{\sigma_X},\frac{Y}{\sigma_Y}\right)$$
$$=1+1-2\rho(X,Y)=0$$

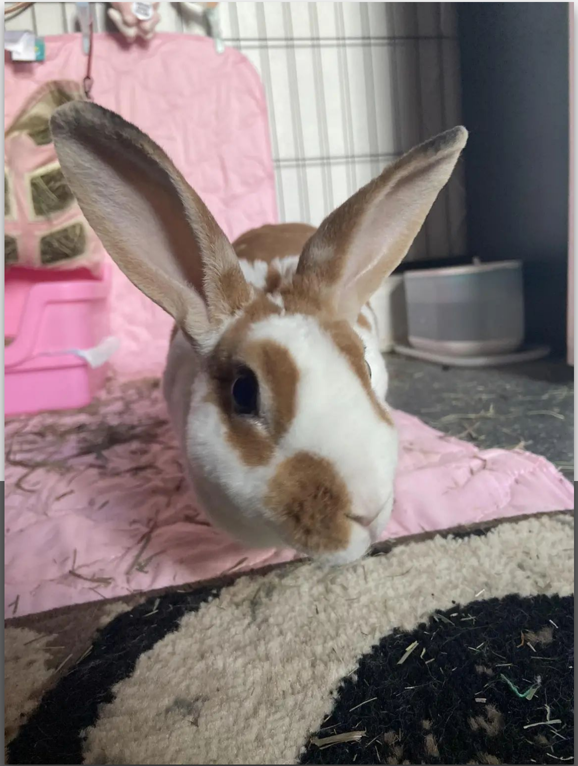
$$\left(\text{Cov}\left(\frac{X}{\sigma_X},\frac{Y}{\sigma_Y}\right)=\frac{1}{\sigma_X}\frac{1}{\sigma_Y}\text{Cov}(X,Y)\right)$$

So $\dfrac{X}{\sigma_X}-\dfrac{Y}{\sigma_Y}$ is a constant with prob. 1

$\Rightarrow$ Y is a linear function of X.

Space for notes 🕛 🕑 🕞

We learned something Today, in Time Series 🕛



What do we tell quin?