STA 542: Introduction to Time Series Analysis

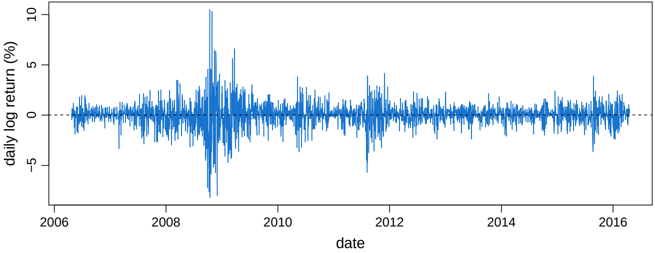
Lecture 6: Estimating uncertainty from one path

Lasse Vuursteen

Duke University

Mon Sep 14, 2026

How can this record supply its own standard error?



Suppose we model these returns by a weakly stationary process with mean μ.

- We estimate μ by the sample mean.
- The population autocovariances are unknown. What goes into its standard error?

Real data: `astsa::djia`, April 2006–April 2016; 2,517 daily log returns in percent.

The variance formula still uses an unknown covariance sum

For a weakly stationary process with ACVF γ, write $v_T := \mathrm{Var}(\bar{X}_T)$ and $q_T := T v_T$. Then

$$q_T = \gamma(0) + 2 \sum_{h=1}^{T-1} \left(1 - \frac{h}{T}\right) \gamma(h).$$

- The weight $1 - h/T$ counts the available pairs.
- Fixed-lag covariance learning controls each fixed h.
- The number of terms here grows with T.

What fixed population quantity could we estimate instead?

Summable covariances give a stable variance scale

Suppose $\sum_{h\in\mathbb{Z}} |\gamma(h)| < \infty$. Define the *long-run variance* by

$$\tau^2 := \gamma(0) + 2\sum_{h=1}^{\infty} \gamma(h).$$

- The scaled variance satisfies $q_T \longrightarrow \tau^2 \geq 0$.
- If $\tau^2 > 0$, then $v_T \approx \tau^2 / T$.

Summable covariances give a stable variance scale

Suppose $\sum_{h\in\mathbb{Z}} |\gamma(h)| < \infty$. Define the *long-run variance* by

$$\tau^2 := \gamma(0) + 2\sum_{h=1}^{\infty} \gamma(h).$$

- The scaled variance satisfies $q_T \longrightarrow \tau^2 \geq 0$.
- If $\tau^2 > 0$, then $v_T \approx \tau^2 / T$.

Each fixed weight tends to one. Why does that alone fail to justify the limit of the whole sum?

First control the tail, then increase the sample size

Keep a cutoff K fixed. For $T > K$, the variance formula gives

$$|q_T - \tau^2| \leq \underbrace{\frac{2}{T}\sum_{h=1}^{K} h|\gamma(h)|}_{\text{fixed finite sum}} + \underbrace{2\sum_{h>K} |\gamma(h)|}_{\text{covariance tail}}.$$

Which term should we control first?

First control the tail, then increase the sample size

Keep a cutoff K fixed. For $T > K$, the variance formula gives

$$|q_T - \tau^2| \le \underbrace{\frac{2}{T}\sum_{h=1}^{K} h|\gamma(h)|}_{\text{fixed finite sum}} + \underbrace{2\sum_{h>K} |\gamma(h)|}_{\text{covariance tail}} .$$

Which term should we control first?

1. Choose K so the covariance tail is small.
2. With that K fixed, increase T so the finite-sum term is small.

Absolute summability supplies the first choice.

Workbook: *The Long-Run Variance.*

Can the long-run variance be zero?

Let $X_t = Z_t - Z_{t-1}$, with centered i.i.d. innovations of variance $\sigma^2 > 0$. Summing the observations cancels the interior terms:

$$\bar{X}_T = \frac{Z_T - Z_0}{T}, \qquad v_T = \frac{2\sigma^2}{T^2}.$$

Can the long-run variance be zero?

Let $X_t = Z_t - Z_{t-1}$, with centered i.i.d. innovations of variance $\sigma^2 > 0$. Summing the observations cancels the interior terms:

$$\bar{X}_T = \frac{Z_T - Z_0}{T}, \qquad v_T = \frac{2\sigma^2}{T^2}.$$

- Here $Tv_T \to 0$, so $\tau^2 = 0$.
- The mean still has positive variance at each finite T.
- The approximation by a positive constant divided by T fails here.

Would a variance estimate of zero be consistent?

Under covariance summability, $v_T \to 0$. Consider

$$\widetilde{v}_T := 0, \qquad |\widetilde{v}_T - v_T| = v_T \longrightarrow 0.$$

Does this make zero an adequate variance estimate?

Would a variance estimate of zero be consistent?

Under covariance summability, $v_T \to 0$. Consider

$$\widetilde{v}_T := 0, \qquad |\widetilde{v}_T - v_T| = v_T \longrightarrow 0.$$

Does this make zero an adequate variance estimate?

Its absolute error vanishes, but its estimated standard error is always zero. We need accuracy relative to the target's size.

Relative accuracy keeps the standard error on scale

Suppose $v_T > 0$ eventually and $\widehat{v}_T \geq 0$. The estimate is *ratio-consistent* if

$$\frac{\widehat{v}_T}{v_T} \xrightarrow{P} 1.$$

- This convergence is in probability, across repeated records.
- It also gives $\sqrt{\widehat{v}_T}/\sqrt{v_T} \xrightarrow{P} 1$.
- The zero estimate has ratio zero whenever $v_T > 0$.

Estimating the stable target gives the required ratio

Suppose $q_T = Tv_T \to \tau^2 > 0$ and $\widehat{\tau}_T^2 \xrightarrow{P} \tau^2$, with $\widehat{\tau}_T^2 \geq 0$. Set $\widehat{v}_T := \widehat{\tau}_T^2 / T$. Then

$$\frac{\widehat{v}_T}{v_T} = \frac{\widehat{\tau}_T^2}{q_T} \xrightarrow{P} \frac{\tau^2}{\tau^2} = 1.$$

Estimating the stable target gives the required ratio

Suppose $q_T = T v_T \to \tau^2 > 0$ and $\widehat{\tau}_T^2 \xrightarrow{P} \tau^2$, with $\widehat{\tau}_T^2 \geq 0$. Set $\widehat{v}_T := \widehat{\tau}_T^2 / T$. Then

$$\frac{\widehat{v}_T}{v_T} = \frac{\widehat{\tau}_T^2}{q_T} \xrightarrow{P} \frac{\tau^2}{\tau^2} = 1.$$

Where did positivity of τ^2 enter the argument?

What happens if we use every available lag?

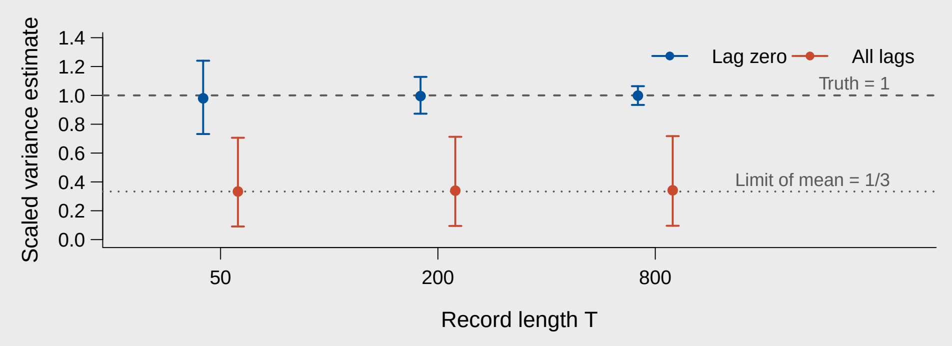
Using the divisor-T sample ACVF $\widehat{\gamma}_T(h)$, consider

$$\widehat{q}_T^{\mathrm{all}} := \widehat{\gamma}_T(0) + 2\sum_{h=1}^{T-1}\left(1 - \frac{h}{T}\right)\widehat{\gamma}_T(h).$$

Suppose the observations are i.i.d. $N(0,1)$, so $q_T = 1$.

Which estimate should improve with T: $\widehat{\gamma}_T(0)$, the all-lag estimate, or both?

Using all lags fails even under independence



Fixed-lag consistency does not control the combined error over all retained lags.

Simulated i.i.d. $N(0,1)$ paths; 2,500 replications per length. Points: means; bars: central 80% ranges.

Workbook: *Why Not Use Every Lag? — bias calculation and proof of inconsistency.*

A known cutoff reduces the task to finitely many lags

Suppose a known integer $m \geq 0$ satisfies $\gamma(h) = 0$ for $|h| > m$. For $T > m$,

$$q_T = \gamma(0) + 2 \sum_{h=1}^{m} \left(1 - \frac{h}{T}\right) \gamma(h).$$

- Only $m + 1$ covariance features need to be learned.
- The number m stays fixed as the record grows.
- A covariance cutoff does not assert independence beyond that lag.

The known-range estimate uses those same lags

Define

$$\widehat{v}_{m,T} := \frac{1}{T} \max\left\{ \widehat{\gamma}_T(0) + 2\sum_{h=1}^{m} \left(1 - \frac{h}{T}\right) \widehat{\gamma}_T(h),\ 0 \right\}.$$

- Retain the exact finite-sample pair weights.
- The positive part prevents a negative variance estimate.

The known-range estimate uses those same lags

Define

$$\widehat{v}_{m,T} := \frac{1}{T} \max\left\{ \widehat{\gamma}_T(0) + 2\sum_{h=1}^{m} \left(1 - \frac{h}{T}\right) \widehat{\gamma}_T(h),\ 0 \right\}.$$

- Retain the exact finite-sample pair weights.
- The positive part prevents a negative variance estimate.

What must the fixed-lag estimates learn for this rule to work?

A fixed number of consistent estimates can be combined

Assume weak stationarity, ACVF learnability through the known cutoff m, and $\tau^2 > 0$.

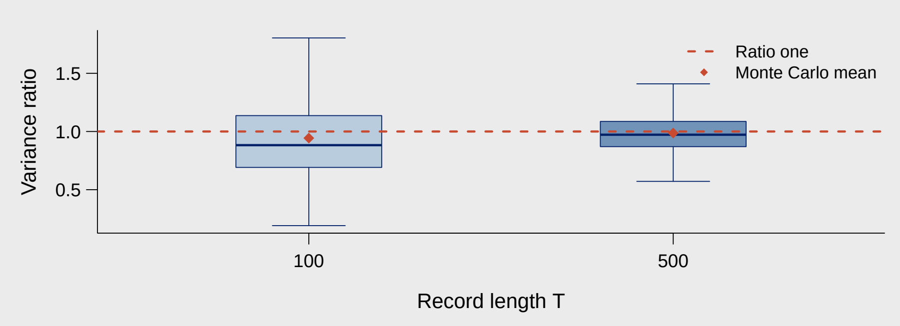
1. Each retained sample covariance converges to its population value.
2. Each weight tends to one; there are only $m + 1$ terms.
3. The positive part has no effect near the positive limit τ^2.

$$\frac{\widehat{v}_{m,T}}{v_T} \xrightarrow{P} 1.$$

After multiplication by T, both quantities converge to the same positive limit.

Longer records improve the estimated variance ratio

For $X_t = \mu + Z_t + 0.5Z_{t-1} + 0.25Z_{t-2}$, take $Z_t = E_t - 1$ with i.i.d. $E_t \sim \text{Exponential}(1)$.



Simulated data: 5,000 stationary paths per length; known range two. Boxplots omit outliers.

Workbook: *Estimating uncertainty for an MA(2) — covariance calculations and R code.*

How many lags should we keep when the range is unknown?

A finite sample ACF cannot certify an exact population cutoff.

Consider a summable ACVF with a nonzero tail.

- What happens if we always retain the same five lags?
- What happens if we retain every available lag?

How many lags should we keep when the range is unknown?

A finite sample ACF cannot certify an exact population cutoff.

Consider a summable ACVF with a nonzero tail.

- What happens if we always retain the same five lags?
- What happens if we retain every available lag?

We need the retained range to grow, while controlling the error from estimating more covariances.

Bartlett weights taper the retained covariances

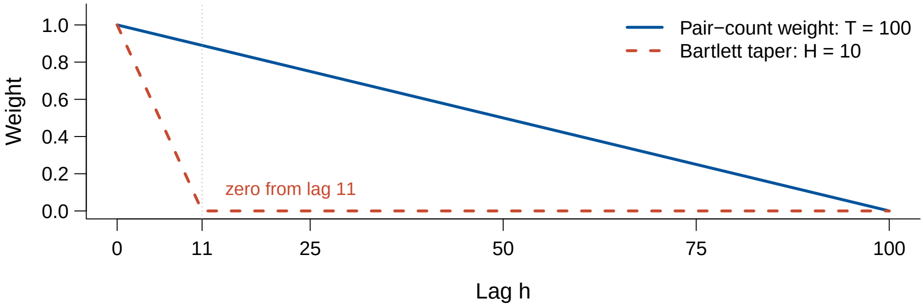
Choose an integer bandwidth $0 \le H < T$. Define

$$\widehat{\tau}^2_{H,T} := \widehat{\gamma}_T(0) + 2\sum_{h=1}^{H}\left(1 - \frac{h}{H+1}\right)\widehat{\gamma}_T(h).$$

- H is the largest retained lag.
- The weights decrease toward the end of the retained range.
- With the divisor-T sample ACVF, this estimate is nonnegative.

The corresponding variance estimate is $\widehat{\tau}^2_{H,T}/T$.

The two triangular weights have different jobs

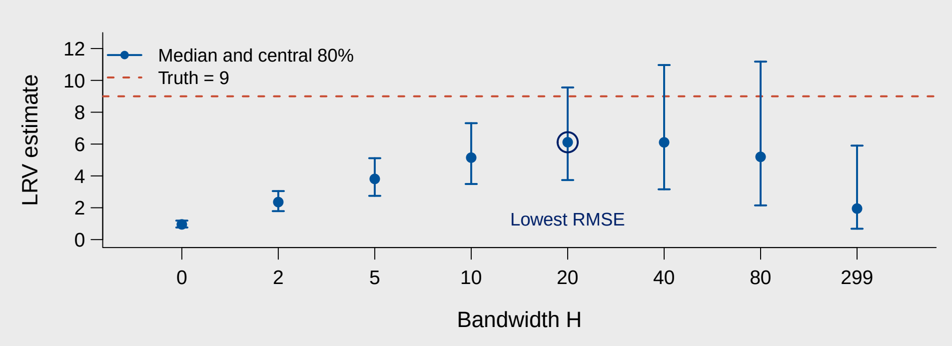


- $1 - h/T$: pair counting in the exact variance formula.
- $1 - h/(H+1)$: a taper chosen for estimation.

Mathematical illustration, $T = 100$ and $H = 10$. At $H = T - 1$, Bartlett reproduces the failed all-lag estimate.

A larger bandwidth can make the estimate more variable

Stationary Gaussian AR(1), $\phi = 0.8$, marginal variance one; the long-run variance is 9.



Which bandwidths trade substantial bias for lower variability?

Simulated data: 3,000 paths, $T = 300$. Points: medians; bars: central 80% ranges.

Workbook: *Bartlett sensitivity for an AR(1) — population targets and RMSE comparison.*

A growing range needs control of the combined error

Assume weak stationarity and $\sum_{h\in\mathbb{Z}} |\gamma(h)| < \infty$. Choose deterministic $H_T \to \infty$ with $H_T < T$.

A sufficient additional condition is

$$\sum_{h=0}^{H_T} |\widehat{\gamma}_T(h) - \gamma(h)| \xrightarrow{P} 0.$$

It controls all the errors in the retained range together.

A growing range needs control of the combined error

Assume weak stationarity and $\sum_{h\in\mathbb{Z}} |\gamma(h)| < \infty$. Choose deterministic $H_T \to \infty$ with $H_T < T$.

A sufficient additional condition is

$$\sum_{h=0}^{H_T} |\widehat{\gamma}_T(h) - \gamma(h)| \xrightarrow{P} 0.$$

It controls all the errors in the retained range together.

Does consistency at every fixed lag establish this condition?

Bartlett consistency requires three errors to vanish

By the triangle inequality,

$$|\widehat{\tau}^2_{H_T,T} - \tau^2| \leq \underbrace{2\sum_{h=0}^{H_T} |\widehat{\gamma}_T(h) - \gamma(h)|}_{\text{estimation error}}$$

$$+ \underbrace{2\sum_{h=1}^{H_T} \frac{h}{H_T+1}|\gamma(h)|}_{\text{taper error}} + \underbrace{2\sum_{h>H_T} |\gamma(h)|}_{\text{omitted tail}}.$$

Combined accuracy handles estimation; summability handles the tail. For the taper, reuse the finite-head / small-tail argument.

Thus $\widehat{\tau}^2_{H_T,T} \xrightarrow{P} \tau^2$.

An error bound can restrict bandwidth growth

Suppose a model supplies $\mathbb{E}[|\widehat{\gamma}_T(h) - \gamma(h)|] \le C/\sqrt{T}$ for all retained lags, with fixed $C < \infty$. Then

$$\mathbb{E}\left[\sum_{h=0}^{H_T} |\widehat{\gamma}_T(h) - \gamma(h)|\right] \le \frac{C(H_T + 1)}{\sqrt{T}}.$$

An error bound can restrict bandwidth growth

Suppose a model supplies $\mathbb{E}[|\widehat{\gamma}_T(h) - \gamma(h)|] \leq C/\sqrt{T}$ for all retained lags, with fixed $C < \infty$. Then

$$\mathbb{E}\left[\sum_{h=0}^{H_T} |\widehat{\gamma}_T(h) - \gamma(h)|\right] \leq \frac{C(H_T + 1)}{\sqrt{T}}.$$

- $H_T \to \infty$ with $H_T = o(\sqrt{T})$ makes this bound vanish.
- Markov's inequality gives combined accuracy in probability.
- The model must supply the bound; this is one sufficient route.

Can an estimated variance calibrate an interval?

Fix a weakly stationary law with mean μ and $v_T > 0$ eventually. Suppose $\widehat{v}_T \geq 0$ and $\widehat{v}_T / v_T \xrightarrow{P} 1$. For $0 < \alpha < 1$, consider

$$C_T := \left[\bar{X}_T - \sqrt{\frac{\widehat{v}_T}{\alpha}}, \quad \bar{X}_T + \sqrt{\frac{\widehat{v}_T}{\alpha}} \right].$$

With known v_T, Chebyshev gives a finite-sample coverage bound.

Here the threshold is random and uses the same path as the mean. What additional argument is needed?

Ratio consistency controls how small the threshold can be

Fix $0 < \varepsilon < 1$. Let $A_T := \{\widehat{v}_T \geq (1-\varepsilon)v_T\}$. Ratio consistency gives $\mathbb{P}(A_T^c) \to 0$.

On A_T, failure of coverage implies

$$|\bar{X}_T - \mu| > \sqrt{\frac{\widehat{v}_T}{\alpha}} \geq \sqrt{\frac{(1-\varepsilon)v_T}{\alpha}}.$$

Ratio consistency controls how small the threshold can be

Fix $0 < \varepsilon < 1$. Let $A_T := \{\widehat{v}_T \geq (1-\varepsilon)v_T\}$. Ratio consistency gives $\mathbb{P}(A_T^c) \to 0$.

On A_T, failure of coverage implies

$$|\bar{X}_T - \mu| > \sqrt{\frac{\widehat{v}_T}{\alpha}} \geq \sqrt{\frac{(1-\varepsilon)v_T}{\alpha}}.$$

The rightmost threshold is deterministic under the fixed law. We can apply Chebyshev.

The failure event splits into two controlled events

Since $\mathbb{E}[\bar{X}_T] = \mu$ and $\mathrm{Var}(\bar{X}_T) = v_T$,

$$\mathbb{P}\{\mu \notin C_T\} \le \mathbb{P}(A_T^c) + \mathbb{P}\left\{ |\bar{X}_T - \mu| > \sqrt{\frac{(1-\varepsilon)v_T}{\alpha}} \right\}$$

$$\le \underbrace{\mathbb{P}(A_T^c)}_{\longrightarrow 0} + \underbrace{\frac{\alpha}{1-\varepsilon}}_{\text{Chebyshev bound}} .$$

The failure event splits into two controlled events

Since $\mathbb{E}[\bar{X}_T] = \mu$ and $\mathrm{Var}(\bar{X}_T) = v_T$,

$$\mathbb{P}\{\mu \notin C_T\} \le \mathbb{P}(A_T^c) + \mathbb{P}\left\{|\bar{X}_T - \mu| > \sqrt{\frac{(1-\varepsilon)v_T}{\alpha}}\right\}$$

$$\le \underbrace{\mathbb{P}(A_T^c)}_{\longrightarrow\ 0} + \underbrace{\frac{\alpha}{1-\varepsilon}}_{\text{Chebyshev bound}} .$$

- First take the limit superior as $T \to \infty$.
- Then let $\varepsilon \downarrow 0$.

The event bound required no independence between the two estimates.

The feasible interval has a limiting lower coverage bound

Under the fixed law and ratio-consistency assumptions just used,

$$\liminf_{T\to\infty} \mathbb{P}\{\mu \in C_T\} \geq 1 - \alpha.$$

- Known-range estimation supplies $\widehat{v}_T$ under its assumptions.
- Bartlett consistency also suffices when $\tau^2 > 0$.

The feasible interval has a limiting lower coverage bound

Under the fixed law and ratio-consistency assumptions just used,

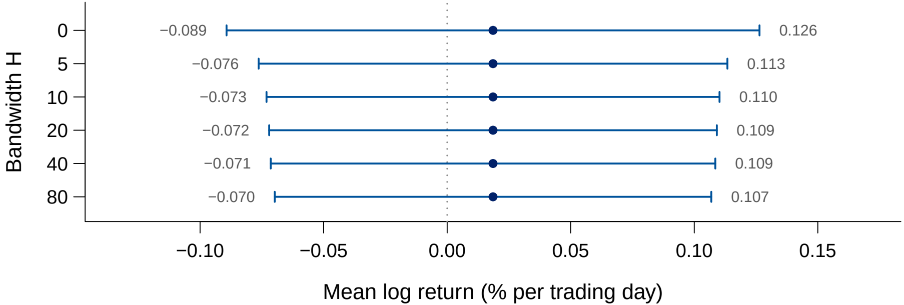
$$\liminf_{T\to\infty} \mathbb{P}\{\mu \in C_T\} \geq 1 - \alpha.$$

- Known-range estimation supplies $\widehat{v}_T$ under its assumptions.
- Bartlett consistency also suffices when $\tau^2 > 0$.

Does the result promise at least 95% coverage for a particular record length? Does it say coverage tends to exactly 95%?

The same return record now gives computable intervals

Compare bandwidths using Bartlett intervals with multiplier $1/\sqrt{0.05}$.



Coverage still requires a law under which the variance estimate is ratio-consistent. These calculations do not establish that assumption.

Real data: `astsa::djia`, April 2006–April 2016; daily log returns in percent.

What would justify a shorter interval?

We have an estimated standard error and a conservative coverage argument.

At 95%, the multiplier is $1/\sqrt{0.05} \approx 4.47$.

- When does a dependent sample mean have an approximately Gaussian shape?
- Would that justify the smaller multiplier 1.96?

Workbook 7 develops the required central limit theorems.