# 5. Sample Autocorrelation, etc.

Thursday, January 20, 2022 1:03 PM

See §1.4.1 and §2.4.2 in the book

## Sample mean, ACVF, ACF

If $\{X_t\}$ is stationary, $\mu := \mathbb{E}[X_t]$, $\sigma^2 = \gamma(0)$

theoretical

$\gamma(h) := \mathrm{Cov}[X_1, X_{1+h}]$

$\rho(h) := \mathrm{Cor}[X_1, X_{1+h}] = \gamma(h)/\gamma(0)$

But given data $\{x_t\}$ (a realization of $\{X_t\}$),
assuming we think $\{X_t\}$ is stationary,
how to find $\mu, \rho, \gamma$? We can't find exactly but we can estimate

**Def** Sample Mean

$$\bar{x} := \frac{1}{n}\sum_{t=1}^{n} x_t$$ sometimes written $\hat{\mu}$ or similar

$$\bar{X} := \frac{1}{n}\sum_{t=1}^{n} X_t$$ } sometimes we use this version ( $\bar{X}$ is estimator, $\bar{x}$ is estimate )

**Fact**: if $\{X_t\}$ really is stationary, then $\bar{X}$ is unbiased,

meaning $\mathbb{E}\,\bar{X} = \mu$. ( proof: $\mathbb{E}[\bar{x}] = \frac{1}{n}\sum_{t=1}^{n}\mathbb{E}[x_t]$ linearity of $\mathbb{E}$

$= \frac{1}{n}\sum_{t=1}^{n}\mu$ stationarity

$= \mu$ )

**Def** Sample autocovariance function (ACVF) and sample autocorrelation function (ACF)

$\widehat{\text{ACVF}}$: $$\hat{\gamma}(h) := \frac{1}{n}\sum_{t=1}^{n-h}(x_{t+h} - \bar{x})(x_t - \bar{x}) \qquad 0 \le h < n$$

and define $\hat{\gamma}(h) = \hat{\gamma}(|h|)$ if $h < 0$

$\widehat{\text{ACF}}$: $$\hat{\rho}(h) := \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)}$$

⚠ Neither $\hat{\gamma}$ nor $\hat{\rho}$ defined for $h \geq n$

⚠ Even if $h < n$, may not be reliable.

( small $h$ = more reliable, since more terms in sum
large $n$ = " " )

eg. $\hat{\rho}(0)$ is always reliable! ( $\hat{\rho}(0) = 1$, and $\rho(0) = 1$ ✓ )

rule-of-thumb (Jenkins '76, p.52 in our text)

$n \geq 50$ and $h \leq n/4$

⚠ It's odd that we sum $n-h$ terms but divide by $n$

(Somewhat analogous to Bessel's correction for $\hat{\sigma}^2$:

$$\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2. \quad \text{Why } \frac{1}{n-1}?$$

a) because $\bar{x}$ is also an estimate, not true mean $\mu$

b) this way it's unbiased, $\mathbb{E}[\hat{\sigma}^2]=\sigma^2$

c) with $1/n$, then if $n=1$, $\hat{\sigma}^2=1$ which is misleading. better to return $\hat{\sigma}^2 = 0/0$ = undefined )

Why include $\frac{1}{n}$ instead of $\frac{1}{n-h}$?

One reason: using $\frac{1}{n}$, then matrices $\hat{\Gamma}_n$ and $\hat{R}_n$ are non-negative definite $\quad (\hat{\Gamma}_n)_{ij} = \hat{\gamma}(i-j) \quad (\hat{R}_n)_{ij} = \hat{\rho}(i-j)$

(a nice property, since population $\Gamma_n$ and $R_n$ are non-neg. def)

However, $\hat{\gamma}$ and $\hat{\rho}$ (using either $\frac{1}{n}$ or $\frac{1}{n-h}$ normalization) are biased, though amount of bias $\to 0$ as $n \to \infty$

## Some properties

### Sample mean $\bar{x}$

- We saw that it's unbiased, $\mathbb{E}[\bar{X}]=\mu$
- Is it consistent? i.e., as $n\to\infty$, does it converge to $\mu$? (has technical meaning; we'll look at sufficient condition)

Yes, under reasonable assumptions:

$\bar{X}_n$ notation denotes we have $n$ samples

Calculate,

$$\mathrm{Var}[\bar{X}_n] = \mathbb{E}(\bar{X}_n-\mu)^2 = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\mathrm{Cov}(X_i,X_j)$$

$$= \frac{1}{n}\sum_{h=-n}^{n}\left(1-\frac{|h|}{n}\right)\cdot\gamma(h)$$

$$\to 0 \text{ if } \underbrace{\gamma(h)\to 0 \text{ as } h\to 0}_{\text{assumption}}$$

See §2.4.1 for details (unimportant)

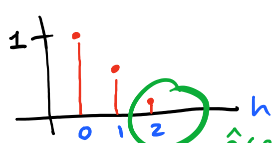
i.e. $\bar{X}_n \xrightarrow{q.m.} \mu$

### Sample ACVF, ACF: $\hat{\gamma}, \hat{\rho}$

motivation: Suppose $\{X_t\} \sim MA(1)$, then $\rho(h)=0$ for $|h|\geq 2$

eg. $\rho(1)=1/2$ if $\theta=1$

We have data $\{x_t\}$, w/ $\hat{\rho}$




$\hat{\rho}(2)$ is small but nonzero

Is this inconsistent with $\rho(2)=0$? (hence ruling out a MA(1) model)?

## Bartlett's Formula

Under regularity conditions (p.52 book gives link to full Thm),

$$\sqrt{n}\,(\hat{\rho}_n - \rho_n) \xrightarrow{d} N(0, W)$$

where $\rho_n = \begin{bmatrix} \rho(1) \\ \vdots \\ \rho(n) \end{bmatrix}$ I'm abusing notation (estimator vs estimate)

i.e. $\hat{\rho}_n \approx N(\rho, \frac{1}{n} W)$

where $W = (w_{ij})$ covariance matrix,

$$w_{ij} = \sum_{k=1}^{\infty} \left(\rho(k+i) + \rho(k-i) - 2\rho(i)\rho(k)\right) \times \left(\rho(k+j) + \rho(k-j) - 2\rho(j)\rho(k)\right)$$

implication...

We don't know $\rho$ so how is formula useful?
For hypothesis tests, for example.

eg., under $H_0$, assume $\{X_t\} \sim WN(0, \sigma^2)$

e.g. residuals after fitting a model

then we do know $\rho$: $\left.\begin{array}{l} \rho(0) = 1 \\ \rho(h) = 0 \quad \text{for } h \neq 0 \end{array}\right\} \rho(h) = \mathbb{1}_{[h=0]}$

so $w_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$ so $W = I_{h \times h}$ ← identity matrix

i.e., $\hat{\rho}(1), \ldots, \hat{\rho}(h)$ are approximately iid $N(0, 1/n)$

So if $\hat{\rho}(h) \sim N(0, 1/n)$
then a standard 2-sided 95% confidence interval (centered at 0, since $\rho(h) = 0$ under $H_0$) is $\pm 1.96 \frac{1}{\sqrt{n}}$

(these are those lines you see in R when using "acf")