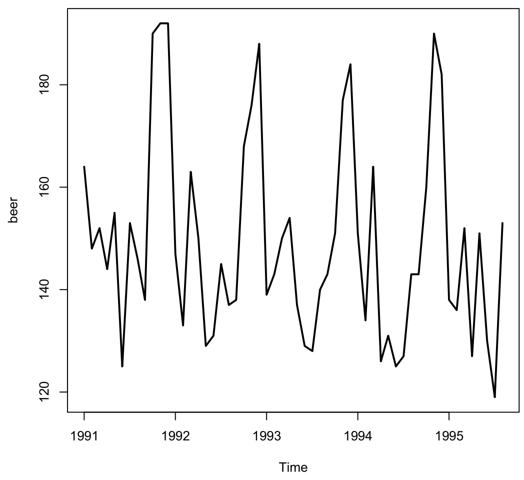
Time series analysis I

Example time series: Beer production. The Monthly Australian beer production has been observed between Jan 1991 to Aug 1995.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1991	164	148	152	144	155	125	153	146	138	190	192	192
1992	147	133	163	150	129	131	145	137	138	168	176	188
1993	139	143	150	154	137	129	128	140	143	151	177	184
1994	151	134	164	126	131	125	127	143	143	160	190	182
1995	138	136	152	127	151	130	119	153				

Time series analysis II



Time series analysis III

Questions:

- What are the patterns in the beer data ?
- What model can you propose to model this stochastic process ? (what explanatory variable would you use ?)

ARIMA models I

Stochastic process s in the time domain

- AR(1)

$$\{s(t+\Delta) = c + \phi_1\ s(t) + \epsilon(t+\Delta)\} \quad \equiv \quad \{s_{t+1} = c + \phi_1\ s_t + \epsilon_{t+1}\}$$

- MA(1)

$$\{s(t+\Delta) = c + \phi_1\ \epsilon(t) + \epsilon(t+\Delta)\} \quad \equiv \quad \{s_{t+1} = c + \phi_1\ \epsilon_t + \epsilon_{t+1}\}$$

with Δ a fixed interval of time.

ARIMA models II

Exercise: Assume a simple AR(1) model:

$$s_t = \alpha\ s_{t-1} + \epsilon_t \quad \text{with } \epsilon_t \sim \mathcal{N}(0, \sigma^2)\ \forall t,\ |\alpha| < 1$$

with observations $\{s_t^{(1)}\}_{t=1,\cdots,n}$. Draw a probabilistic graph to model this time series.

Gaussian Markov Random Fields - Theory and Applications , H. Rue & L. Held, 2005 (chp. 1).

ARIMA models III

Exercise: Assume a simple AR(1) model:

$$s_t = \alpha\ s_{t-1} + \epsilon_t \quad \text{with } \epsilon_t \sim \mathcal{N}(0, \sigma^2)\ \forall t,\ |\alpha| < 1$$

Show

$$p(s_t|s_{t-1}, \cdots, s_1) = p(s_t|s_{t-1}) = \mathcal{N}(s_t; \alpha\ s_{t-1}, \sigma^2)$$

ARIMA models IV

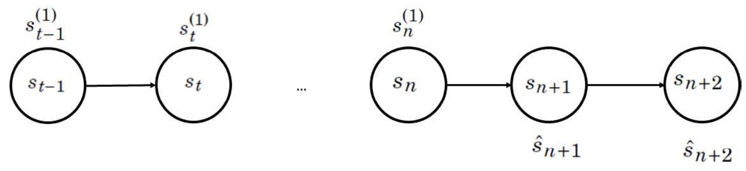
Exercise: Given a temporal stochastic process s with the following Markov property

$$p(s_t|s_{t-1}) = \mathcal{N}(s_t; \alpha\ s_{t-1}, \sigma^2) \quad \forall t$$

with observations $\{s_t^{(1)}\}_{t=1,\cdots,n}$.

- Estimate s at a new time $n+1$ with its confidence interval.
- Give an estimate for s at a new time $n+2$ with its confidence interval.

ARIMA models V



ARIMA models VI

We know

$$p(s_t|s_{t-1}) = \mathcal{N}(s_t; \alpha\ s_{t-1}, \sigma^2) \quad \forall t$$

so at $n+1$:

$$p(s_{n+1}|s_n) = \mathcal{N}(s_{n+1}; \alpha\ s_n, \sigma^2)$$

and we use the observation $s_n^{(1)}$ that is available for s_n:

$$p(s_{n+1}|s_n^{(1)}) = \mathcal{N}(s_{n+1}; \alpha\ s_n^{(1)}, \sigma^2)$$

A good estimate is the expectation $\hat{s}_{n+1} = \alpha\ s_n^{(1)}$ and the 95% prediction interval is $\hat{s}_{n+1} \pm 2\sigma$.

ARIMA models VII

At $n+2$:

$$p(s_{n+2}|s_{n+1}) = \mathcal{N}(s_{n+2}; \alpha\ s_{n+1}, \sigma^2)$$

But we dont have observation for s_{n+1} ! so instead we use:

$$\begin{aligned}
p(s_{n+2}|s_n) &= \int p(s_{n+2}, s_{n+1}|s_n)\ ds_{n+1} \\
&= \int p(s_{n+2}|s_{n+1}, s_n)\ p(s_{n+1}|s_n)\ ds_{n+1} \\
&= \int p(s_{n+2}|s_{n+1})\ p(s_{n+1}|s_n)\ ds_{n+1} \\
&= \int \mathcal{N}(s_{n+2}; \alpha\ s_{n+1}, \sigma^2)\ \mathcal{N}(s_{n+1}; \alpha\ s_n, \sigma^2)\ ds_{n+1} \\
&= \mathcal{N}(s_{n+2}; \alpha^2\ s_n, \sigma^2\ (1+\alpha^2))
\end{aligned}$$

A good estimate is the expectation $\hat{s}_{n+2} = \alpha^2\ s_n^{(1)}$ using the observation $s_n^{(1)}$, and the 95% prediction interval is $\hat{s}_{n+2} \pm 2\sigma\ \sqrt{1+\alpha^2}$.

ARIMA models VIII