时间序列分析（TIME SERIES ANALYSIS）

主讲：吴尚
复旦大学管理学院统计与数据科学系

ARMA模型预测

- 条件数学期望
- 最小均方误差预测
- ARIMA模型预测
- 区间估计
- ARIMA预测的更新
- 对数变换的预测

条件数学期望

- X,Y皆为离散型随机变量, 则Y 对于给定$X=x$的条件数学期望定义为: $E(Y|X=x)=\sum_y y\cdot p_{Y|X}(y|x)$
- 对于一般的函数$h(x)$有:

$$E(h(Y)|X=x)=\sum_y h(y)\cdot p_{Y|X}(y|x)$$

- 特别的:

$$Var(Y|X=x)=\sum_y \big(y-E(Y|X=x)\big)^2\cdot p_{Y|X}(y|x)$$
$$=E(Y^2|X=x)-\big(E(Y|X=x)\big)^2$$

条件数学期望

- X,Y皆为连续型随机变量, 则Y 对于给定$X=x$的条件数学期望定义为：$E(Y|X=x)=\int_{-\infty}^{\infty} y\cdot f_{Y|X}(y|x)dy$
- 对于一般的函数$h(x)$有：

$$E(h(Y)|X=x)=\int_{-\infty}^{\infty} h(y)\cdot f_{Y|X}(y|x)\,dy$$

- 特别的：

$$Var(Y|X=x)=\int_{-\infty}^{\infty}[y-E(Y|X=x)]^2\cdot f_{Y|X}(y|x)\,dy$$

$$=E(Y^2|X=x)-\big(E(Y|X=x)\big)^2$$

例

设随机向量(X,Y)的联合密度为:

$$f_{X,Y}(x,y)=\begin{cases}6xy(2-x-y), & 0\leq x\leq 1, 0\leq y\leq 1\\ 0, & \text{其它}\end{cases}$$

求：$E[Y|X=x], Var(Y|X=x)$

条件数学期望的两重性

(1) $E[Y|X] = g(X)$

其中，$g(x) = E[Y|X = x]$

(2)$E[Y|X_1, \cdots, X_n] = g(X_1, \cdots, X_n)$,

其中，$g(x_1, \cdots, x_n) = E[Y|X_1 = x_1, \cdots, X_n = x_n]$

条件数学期望的性质

- 全期望公式： $E[E[Y|X]] = E[Y]$

$$E[E[Y|X_1,\cdots,X_n]] = E[Y]$$

- 线性： $E[1|Y_1,\cdots,Y_n] = 1$，对任意常数a,b，有

$$E[aY_1 + bY_2|X_1,\cdots,X_n]$$
$$= aE[Y_1|X_1,\cdots,X_n] + bE[Y_2|X_1,\cdots,X_n]$$

- 独立公式：如果Y与$X_1,\cdots,X_n$相互独立，则

$$E[Y|X_1,\cdots,X_n] = E[Y]$$

- ?分解公式：对任意n元连续函数f，有

$$E[f(X_1,\cdots,X_n)Y|X_1,\cdots,X_n]$$
$$= f(X_1,\cdots,X_n)E[Y|X_1,\cdots,X_n]$$

最小均方误差预测

- 我们的目标是用X来预测Y，标准为最小化均方误差，即需要选择一个函数$h(X)$，使得下式达到最小：
$$E[Y-h(X)]^2$$
- 不难证明，最小均方误差预测为
$$h(X)=E[Y|X]$$
- 同理，如果用$X_1,\cdots,X_n$来预测Y，最小均方误差预测为
$$h(X_1,\cdots,X_n)=E[Y|X_1,\cdots,X_n]$$

时间序列的预测

- 假设我们已知序列$Y_1, \cdots, Y_t$，预测未来l期的值Y_{t+l}，则最小均方误差预测记为
$$\hat{Y}_t(l) = E(Y_{t+l} | Y_1, \cdots, Y_t)$$
- 预测误差记为
$$e_t(l) = Y_{t+l} - \hat{Y}_t(l)$$
- 若$E[e_t(l)] = 0$，则称预测是无偏的。
- 预测误差的方差为$Var(e_t(l))$

ARIMA模型预测

- 对于可逆模型，当$j \le 0$时，$E\left[e_{t+j} \middle| Y_1, \cdots, Y_t\right] \approx e_{t+j}$
- ARIMA模型表达式两边同时对$Y_1, \cdots, Y_t$求条件期望得预测递推式：

$$\hat{Y}_t(l) = \phi_1 \hat{Y}_t(l-1) + \cdots + \phi_p \hat{Y}_t(l-p) + \theta_0 - \theta_1 e_{t+l-1}^* - \cdots - \theta_q e_{t+l-q}^*$$

- 其中$e_{t+j}^* = \begin{cases} e_{t+j} & j \le 0 \\ 0 & j > 0 \end{cases}$
- 用模型表达式减去预测递推式可得误差递推式

$$e_t(l) = \phi_1 e_t(l-1) + \cdots + \phi_p e_t(l-p) + e_{t+l} - \theta_1 e'_{t+l-1} - \cdots - \theta_q e'_{t+l-q}$$

- 其中$e'_{t+j} = \begin{cases} 0 & j \le 0 \\ e_{t+j} & j > 0 \end{cases}$

ARIMA模型预测

- 当$l>q$时，预测递推式为非齐次线性差分方程：

$$\hat{Y}_t(l)=\phi_1\hat{Y}_t(l-1)+\cdots+\phi_p\hat{Y}_t(l-p)+\theta_0$$

- 回顾讲义《线性差分方程》
- 误差递推式

$$e_t(l)=\phi_1 e_t(l-1)+\cdots+\phi_p e_t(l-p)+e_{t+l}-\theta_1 e'_{t+l-1}-\cdots-\theta_q e'_{t+l-q}$$

可以简写为

$$\Phi(B)e_t(l)=\Theta(B)e'_{t+l}$$

于是

$$e_t(l)=\Psi(B)e'_{t+l}=e_{t+l}+\psi_1 e_{t+l-1}+\cdots+\psi_{l-1}e_{t+1}$$

- 误差方差为$Var(e_t(l))=\sigma_e^2(1+\psi_1^2+\cdots+\psi_{l-1}^2)$

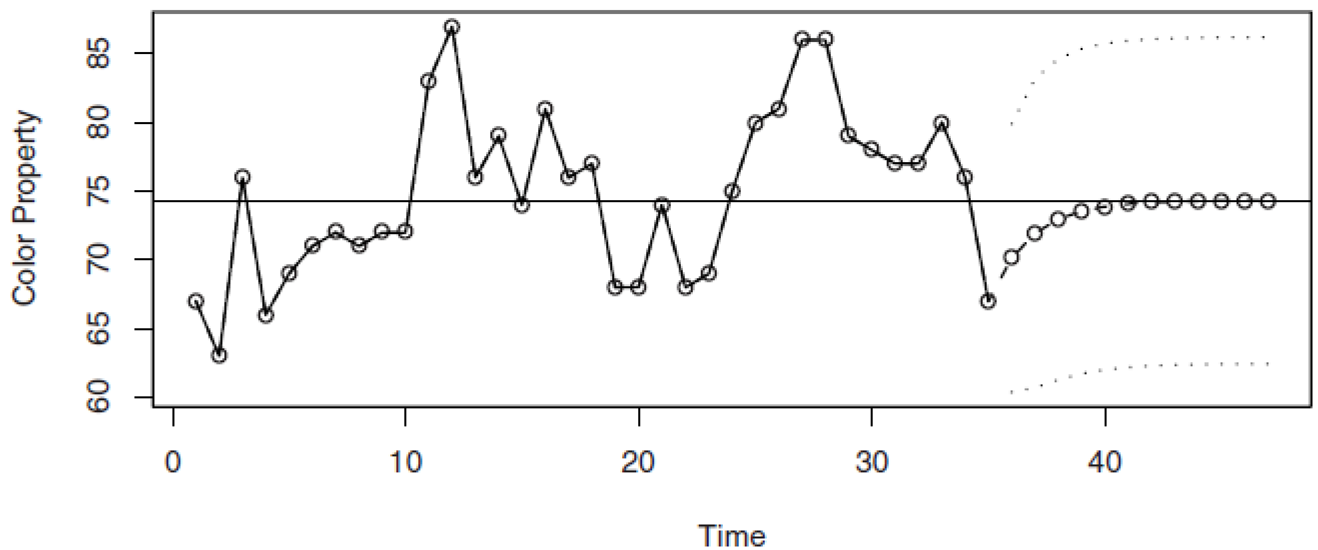
区间估计

- 已知序列$Y_1,\cdots,Y_t$，未来真实值Y_{t+l}的均值为$\hat{Y}_t(l)$，方差为$Var(e_t(l)) = \sigma_e^2(1 + \psi_1^2 + \cdots + \psi_{l-1}^2)$
- 可以认为$\frac{Y_{t+l}-\hat{Y}_t(l)}{\sqrt{Var(e_t(l))}}$大致服从标准正态分布
- $P\left(-z_{1-\frac{\alpha}{2}} < \frac{Y_{t+l}-\hat{Y}_t(l)}{\sqrt{Var(e_t(l))}} < z_{1-\frac{\alpha}{2}}\right) \approx 1 - \alpha$
- 区间估计为

$$\left(\hat{Y}_t(l) - z_{1-\frac{\alpha}{2}}\sqrt{Var(e_t(l))},\ \hat{Y}_t(l) + z_{1-\frac{\alpha}{2}}\sqrt{Var(e_t(l))}\right)$$

例

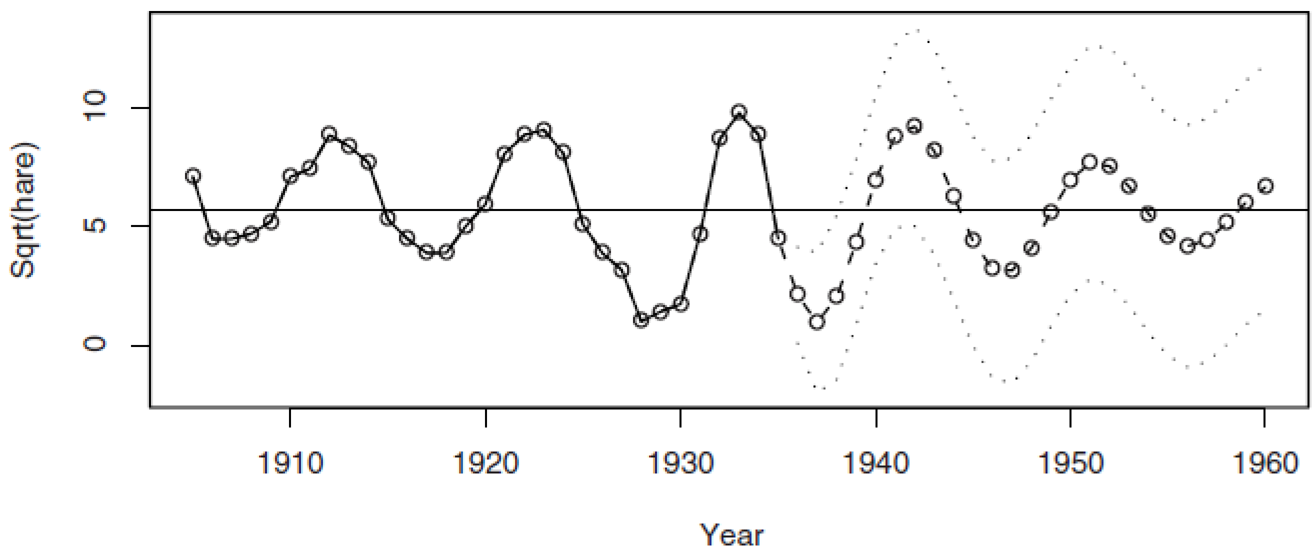
Exhibit 9.3 Forecasts and Forecast Limits for the AR(1) Model for Color




```
> data(color)
> m1.color=arima(color,order=c(1,0,0))
> plot(m1.color,n.ahead=12,type='b',xlab='Time',
   ylab='Color Property')
> abline(h=coef(m1.color)[names(coef(m1.color))=='intercept'])
```

例

Exhibit 9.4 Forecasts from an AR(3) Model for Sqrt(Hare)




```
> data(hare)
> m1.hare=arima(sqrt(hare),order=c(3,0,0))
> plot(m1.hare, n.ahead=25,type='b',
   xlab='Year',ylab='Sqrt(hare)')
> abline(h=coef(m1.hare)[names(coef(m1.hare))=='intercept'])
```

ARIMA预测的更新

- ARIMA模型的预测本质上就是把Y_{t+l}表示成$C_t(l)+I_t(l)$的形式，其中$C_t(l)$是$Y_t, Y_{t-1}, \dots$（和$e_t, e_{t-1}, \dots$）的某个函数，$I_t(l)$是新息项$e_{t+1}, e_{t+2}, \dots, e_{t+l}$的函数，事实上，
$$I_t(l) = e_{t+l} + \psi_1 e_{t+l-1} + \cdots + \psi_{l-1} e_{t+1}$$
- 对于可逆模型，当t充分大时，$\hat{Y}_t(l) = E(Y_{t+l} | Y_1, \cdots, Y_t) \approx C_t(l)$，$e_t(l) \approx I_t(l)$
- 由于$Y_{t+l+1} = C_t(l+1) + e_{t+l+1} + \psi_1 e_{t+l} + \cdots + \psi_l e_{t+1}$，可以发现

$$\hat{Y}_{t+1}(l) = C_{t+1}(l) = C_t(l+1) + \psi_l e_{t+1} = \hat{Y}_t(l+1) + \psi_l [Y_{t+1} - \hat{Y}_t(1)]$$

对数变换的预测

- 给定$Y_t, Y_{t-1}, \dots, Y_1$之后，Y_{t+l}可以表示为确定的常数$\hat{Y}_t(l)$加上随机项$e_t(l)$，由于$e_t(l)$与$Y_t, Y_{t-1}, \dots, Y_1$独立，所以

$$E(Y_{t+l}|Y_1, \cdots, Y_t) = \hat{Y}_t(l)$$
$$Var(Y_{t+l}|Y_1, \cdots, Y_t) = Var(e_t(l))$$

- 对于正态误差，如果$X_t = \exp(Y_t)$，则

$$E(X_{t+l}|X_1, \cdots, X_t) = \exp\left[\hat{Y}_t(l) + \frac{Var(e_t(l))}{2}\right]$$

- 如果最优预测是在给定$X_1, \cdots, X_t$下X_{t+l}分布的中位数，则该预测为$\exp\left[\hat{Y}_t(l)\right]$.