Statistics 531/Econ 677
Midterm, Winter 2008

Let x_t be annual flow of the river Nile, measured at Aswan, between 1871 and 1970. x_t is graphed in Fig. 1.

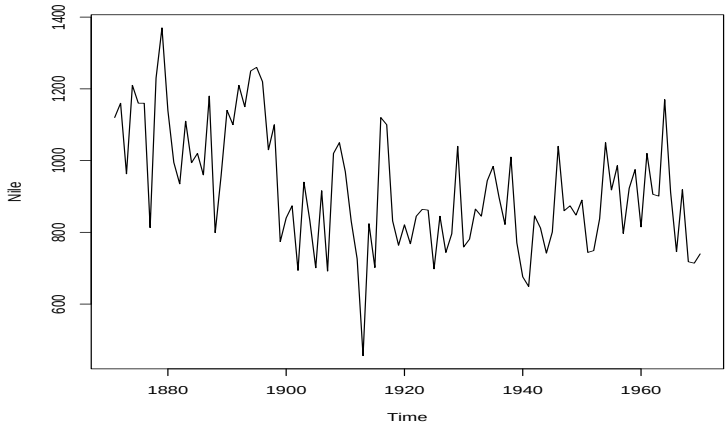


Figure 1: Annual flow of the river Nile, measured at Aswan

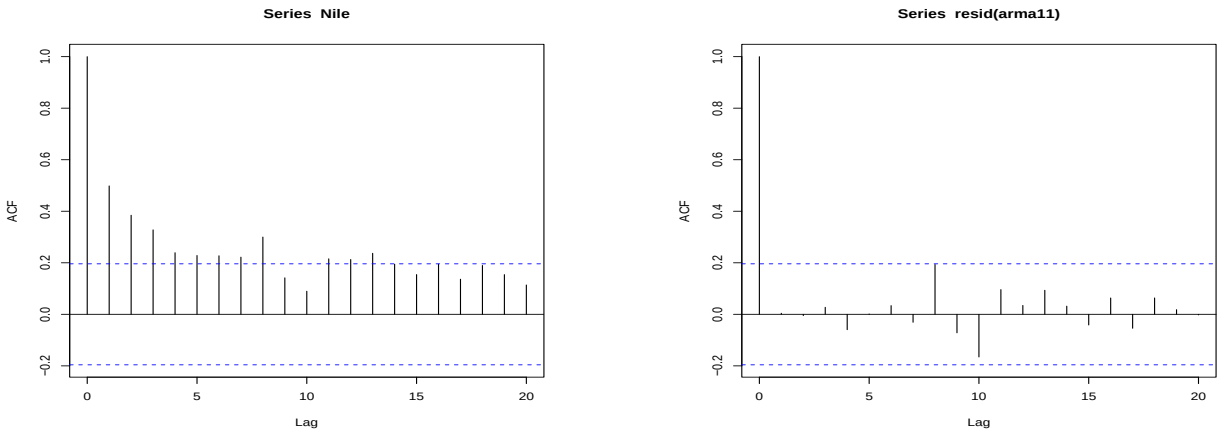
Section A [7 points]. A linear regression model $x_t = \beta_0 + \beta_1 t + \epsilon_t$ was fitted using least squares, via the R command `lm(Nile~time(Nile))`. The output is

```
summary(lm(Nile~time(Nile)))
Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 6132.1736  1001.7578   6.121 1.92e-08 ***
time(Nile)    -2.7143     0.5216  -5.204 1.07e-06 ***
```

[A1, 1 pt]. What is the name of the time series model which makes the assumptions implicit in the above analysis?
[A2, 2 pts]. The estimate $\hat{\beta}_1 = -2.71$ is (a) too high; (b) too low; (c) about right. Choose (a), (b) or (c) and explain briefly.
[A3, 2 pts]. The standard error $SE(\hat{\beta}_1) = 0.53$ is (a) too high; (b) too low; (c) about right. Choose (a), (b) or (c) and explain briefly.
[A4, 2 pts]. If possible from the information given, explain how to make an appropriate test of the hypothesis that $\beta_1 = 0$. If it is not possible, explain why.

Section B [9 points]. An ARMA(1,1) model was fitted using the R command `arima(Nile, order = c(1,0,1))`. The output is given below. The sample ACF of x_t is shown in Fig. 2(a).

```
arima(Nile,order=c(1,0,1))
Coefficients:
         ar1      ma1  intercept
      0.8611  -0.5177   920.5567
```



(a) **(b)**

Figure 2: **(a)** Sample ACF for x_t. **(b)** Sample ACF for residuals from fitting ARMA(1,1) to x_t.

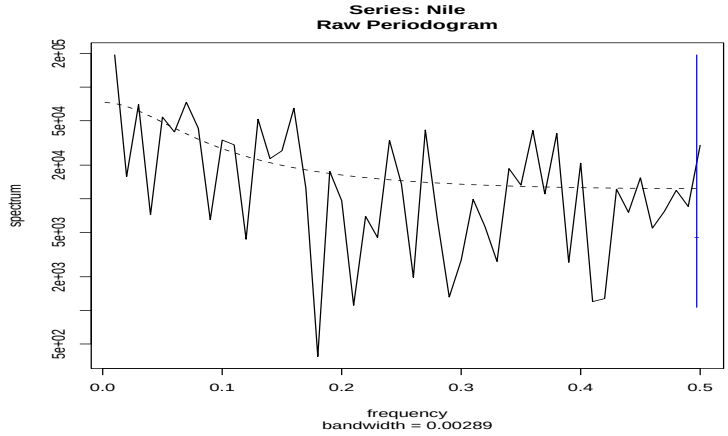


Figure 3: Periodogram, calculated by `spectrum(Nile)`, with the fitted ARMA(1,1) spectrum

```
s.e.  0.1067    0.1908     46.6736

sigma^2 estimated as 19892:  log likelihood = -637.04,  aic = 1282.08
```

[B1, 3 pts]. Write out the fitted model, being carful to specify all model assumptions.
[B2, 2 pts]. The sample ACF of the residuals was plotted by `acf(resid(arma11))`. (see Fig. 2(b)). Explain how the residuals are defined.
[B3, 2 pts]. Explain carefully what the dashed lines show in Fig. 2(a) and 2(b). Discuss briefly how these lines help us to understand the data.
[B4, 2 pts]. AR(1) and AR(2) models both have noticeably large AIC than ARMA(1,1). Explain why Fig. 2(a) suggests that an MA model is not suitable, and argue that ARMA(1,1) is a reasonable model for these data.

Section C [6 pts]. The raw periodogram of x_t is shown in Fig. 3 with the spectral density of the fitted ARMA(1,1) model shown dashed.

[C1, 3 pts]. Give a formula that could be used to calculate a raw periodogram. Explain briefly how (if at all) the solid line in Fig. 3 is calculated differently from your formula.
[C2, 3 pts]. An interesting question is whether the Nile data have a trend or can be modeled as stationary. Discuss this point. There is not necessarily a correct answer, but you should comment on relevant aspects of Sections A, B and C.