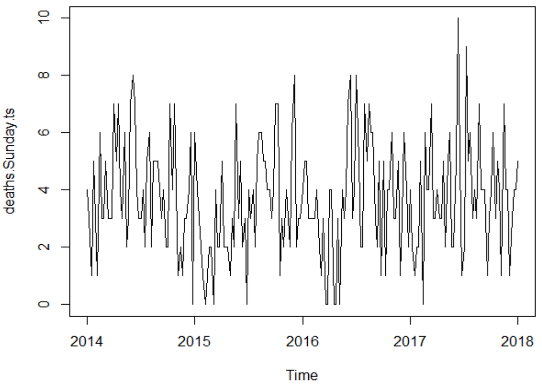
Practice 4. Discrete valued time series

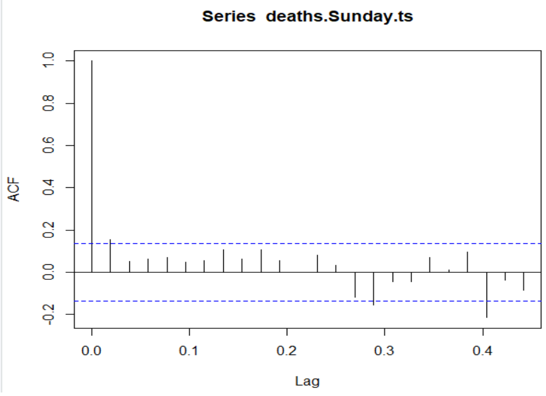
Joan Puigdomenech i Joel Romia

1. Plot the time series, its ACF and its PACF.

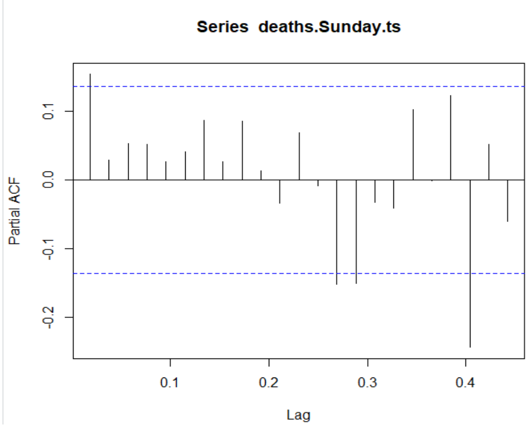
```
> load("./AccidSunday.Rdata")
> plot(deaths.Sunday.ts, type = 'l')
```




```
> acf(deaths.Sunday.ts)
```

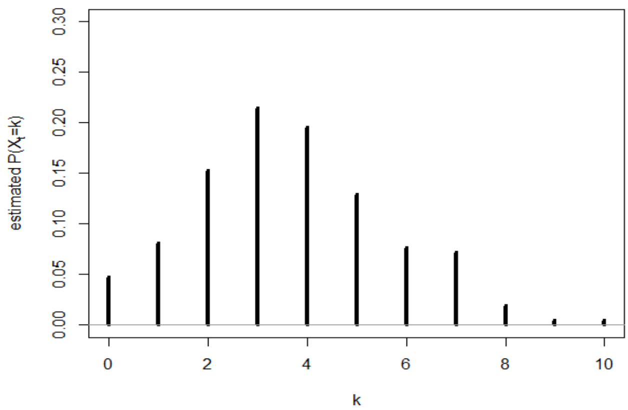



```
> pacf(deaths.Sunday.ts)
```



2. Plot the relative frequencies of the marginal distribution of the time series.

```
> absfreq <- tabulate(deaths.Sunday.ts+1) #+1 to include 0
> Tlen <- length(deaths.Sunday.ts)
> maxval <- max(deaths.Sunday.ts)
> plot(0:maxval, absfreq/Tlen, type="h", xlab = "k", ylab =
expression(paste("estimated P(X"[t],"=k)")), lwd=4, ylim=c(0,0.3))
> abline(h=0,col=8
```



3. Compute the observations' mean $\hat{\mu}=\bar{x}$ and the observations' variance: $\hat{\sigma}^2=\hat{\gamma}_0$. Assume now that the time series follows an INAR(1) model. Then estimate the mean and the variance of the innovations time series: $\hat{\mu}_\varepsilon=\hat{\mu}(1-\hat{\alpha})$, $\hat{\sigma}^2=(1-\hat{\alpha}^2)\hat{\sigma}^2-\hat{\alpha}\hat{\mu}_\varepsilon$.

```
> mn <- mean(deaths.Sunday.ts)
> mn
[1] 3.669856
> varc <- var(deaths.Sunday.ts)
> varc
[1] 3.91452
>
> rho1 <- acf(deaths.Sunday.ts, plot=FALSE)[[1]][2]
> rho1
[1] 0.1532521
> mue <- mn*(1-rho1)
> mue
[1] 3.107443
>
> se <- (1-rho1^2)*varc-rho1*mue
```

```
> se
[1] 3.34636
```

4. To test the equidispersion for the time series, compute the observed Index of Dispersion (ID), $\hat{I}=\hat{\gamma}_0/\bar{X}$, and the 95% confidence interval for the ID. Does 1 belong to this interval? Can we reject the null hypothesis of equidispersion?

```
> ID <- (Tlen-1)/Tlen*varc/mn
> ID
[1] 1.061565
>
> sdID <- sqrt(2/Tlen*(1+rho1^2)/(1-rho1^2))
> ID + -1/Tlen*(1+rho1)/(1-rho1) + c(-1,1)*qnorm(0.975)*sdID
[1] 0.8587609 1.2513351
```

El numero 1 si que pertenece a este rango y por lo tanto no podemos descartar la hipotesis nula de equidispersion.

5. Fit the INAR(1) model to the time series:

- **Maximum likelihood estimation of the Poisson INAR(1) model?**

- **Estimates: $\hat{\lambda},\hat{\alpha}$?**

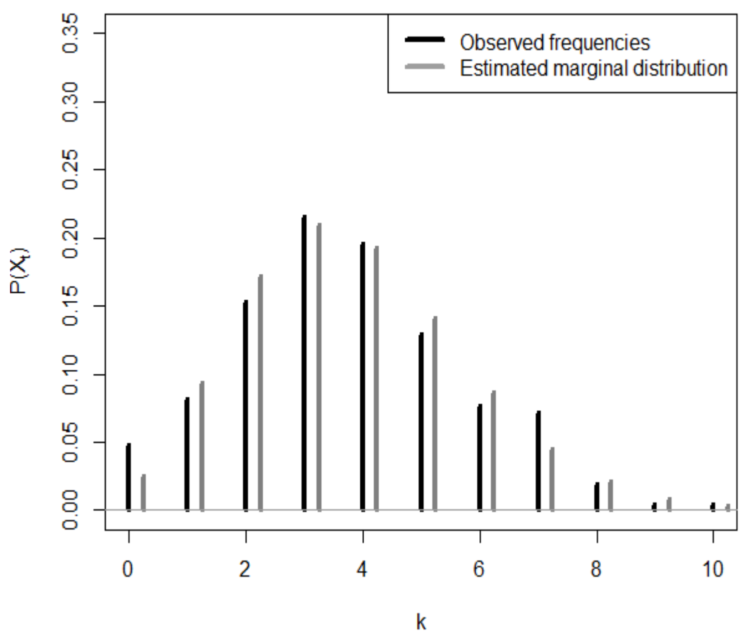
```
> estimate <- c(lambdaestml,alphaestml)
> estimate
[1] 3.1462443 0.1429882
```

- **Observations' Poisson parameter?**

```
> muestml <- lambdaestml/(1-alphaestml)

> muestml
[1] 3.67118
```

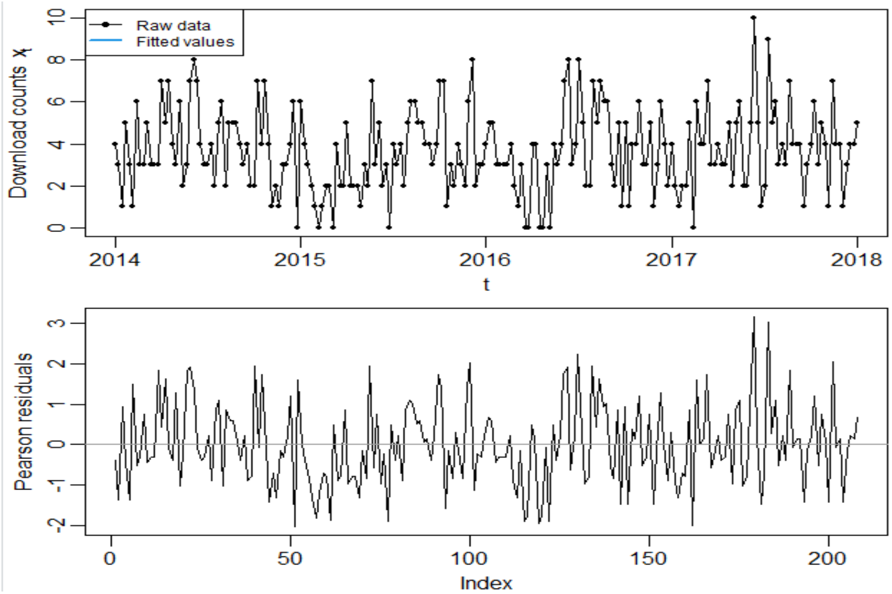
6. Compare the observed relative frequencies with the corresponding probabilities of a Poisson distribution with parameter equal to the estimated one.



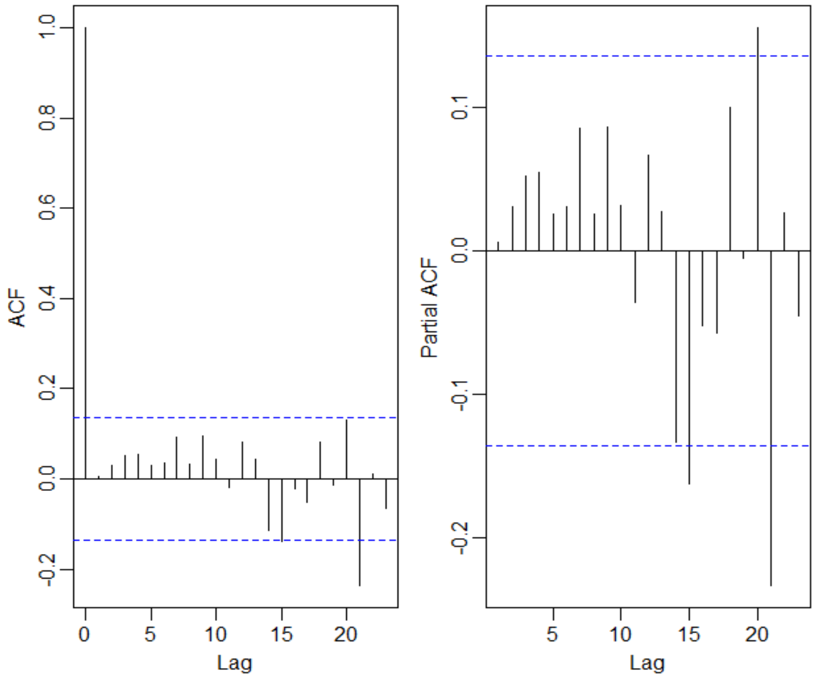
7. Compute the Pearson residuals:

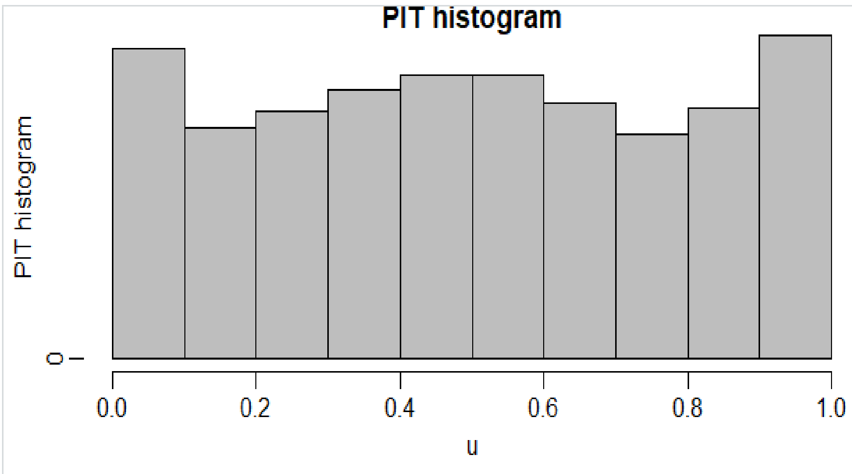
$$e_t=\frac{x_t-E(X_t|X_{t-1})}{\sqrt{Var(X_t|X_{t-1})}}, \quad t=2,\dots,T.$$

Plot the fitted values over the observed time series. In a different graphc, plot the Pearson residuals.



8. Plot the ACF and the PACF of the Pearson residuals. Plot the Probability integral transform (PIT). Do you think the the INAR(1) model is appropriate for this time series?





9. **Fit a Generalized Linear Models for the time series using the function tsglm. Use lags 1 and 52 for the observations and for the mean. Use the logarithm as the link function.**

```
> library(tscount)
> campyfit_pois <- tsglm(campy, model=list(past_obs=c(1,52)), link="log",
distr="poisson")
> summary(campyfit_pois)

Call:
tsglm(ts = campy, model = list(past_obs = c(1, 52)), link = "log",
    distr = "poisson")

Coefficients:
             Estimate  Std.Error  CI(lower)  CI(upper)
(Intercept) 0.645635     0.1905      0.272      1.019
beta_1      0.721458     0.0487      0.626      0.817
beta_52     0.000425     0.0749     -0.146      0.147
Standard errors and confidence intervals (level =  95 %) obtained
by normal approximation.

Link function: log
Distribution family: poisson
Number of coefficients: 3
Log-likelihood: -438.8885
AIC: 883.7769
BIC: 892.6019
QIC: 883.7769
```

10. **Fit now the log-linear model until the middle of 2017 (end=c(2017,26)). Then predict the number of deaths by accidents in Sundays from the middle of 2017 to the end of 2017. Finally, compare the model predictions with the true observed values.**

```
> campyfit_pois_17 <- tsglm(campy_until_2017,
+                           model=list(past_obs=c(1,13)),
+                           link="log", distr="poisson")
> summary(campyfit_pois_17)

Call:
tsglm(ts = campy_until_2017, model = list(past_obs = c(1, 13)),
    link = "log", distr = "poisson")

Coefficients:
             Estimate  Std.Error  CI(lower)  CI(upper)
(Intercept)    0.377     0.1406      0.101      0.653
beta_1         0.597     0.0564      0.487      0.708
beta_13        0.237     0.0570      0.125      0.349
```

```
Standard errors and confidence intervals (level =  95 %) obtained
by normal approximation.

Link function: log
Distribution family: poisson
Number of coefficients: 3
Log-likelihood: -429.0057
AIC: 864.0115
BIC: 872.8364
QIC: 864.0115
```