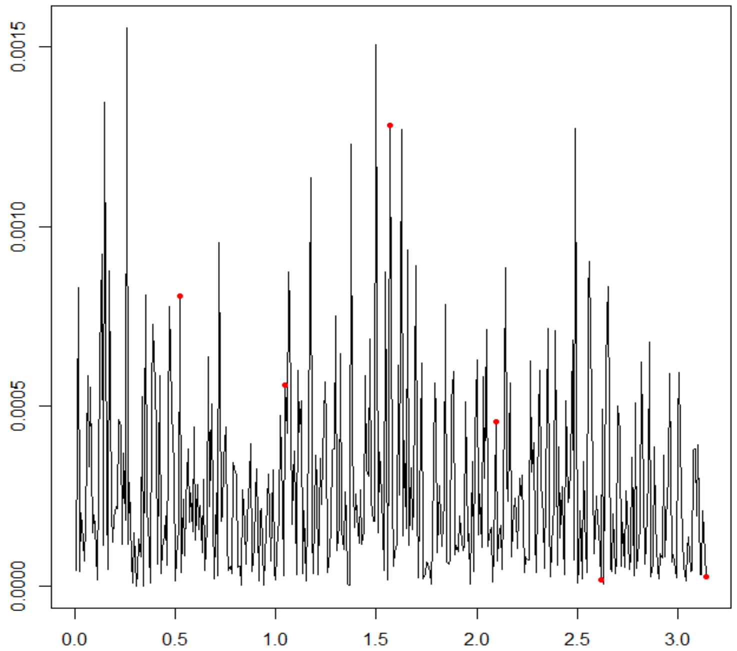
Exercise: Get the monthly S&P 500 index (**^GSPC**) and plot the periodogram of the returns. Make sure that the number of returns is divisible by 12.

```
Y <- read.csv("^GSPC.csv",na.strings="null")
Y <- na.omit(Y) # rows with missing values are omitted
N <- nrow(Y) # number of rows
d <- as.Date(Y[,1]) # dates in column 1
y <- log(Y[,6]) # adjusted close prices in column 6
r <- y[2:N]-y[1:(N-1)] ; n <- N-1    # log returns

n12 <- n-n%%12   # 29%%12=29 modulo 12=5
r12 <- r[(n-n12+1):n]   # r12 contains the last n12 obs.
m12 <- floor(n12/2); f12 <- (2*pi/n12)*(1:m12)
ft12 <- fft(r12)   # Fourier transform
ft12 <- ft12[2:(m12+1)]   # only frequencies 1,…,m
pg12 <- (1/(2*pi*n12))*(Mod(ft12))^2   # periodogram
par(mar=c(2,2,1,1)); plot(f12,pg12,type="l")

n.years <- n12/12  # number of full years
s <- n.years*(1:6)  # indices of seasonal frequencies
lines(f12[s],pg12[s],type="p",pch=20,col="red")
```



Since the periodogram ordinates at the first four seasonal frequencies are relatively large, the returns could possibly contain a small seasonal component.

Exercise: Show that AC

$$\mathrm{Cov}(\alpha x+\beta y,\gamma u+\delta v) = \alpha\gamma\mathrm{Cov}(x,u) + \alpha\delta\mathrm{Cov}(x,v) + \beta\gamma\mathrm{Cov}(y,u) + \beta\delta\mathrm{Cov}(y,v).$$

Suppose that x is white noise with mean μ and variance σ^2 and $\omega_k = \frac{2\pi\cdot k}{n}$, $0<k<\frac{n}{2}$. Then

$$E\hat{A}_k = E\,\tfrac{2}{n}\sum_{t=1}^{n} x_t\cos(\omega_k t)$$

$$= \tfrac{2}{n}\sum_{t=1}^{n}\underbrace{E(x_t)}_{=\mu}\cos(\omega_k t)$$

$$= \tfrac{2\mu}{n}\underbrace{\sum_{t=1}^{n}\cos(\omega_k t)}_{=0} = 0,$$

$$E\hat{B}_k = E\,\tfrac{2}{n}\sum_{t=1}^{n} x_t\sin(\omega_k t) = \ldots = \tfrac{2\mu}{n}\underbrace{\sum_{t=1}^{n}\sin(\omega_k t)}_{=0} = 0,$$

$$\mathrm{Var}(\hat{A}_k) = \tfrac{4}{n^2}\sum_{t=1}^{n}\underbrace{Var(x_t)}_{=\sigma^2}\cos^2(\omega_k t)$$

$$= \tfrac{4\sigma^2}{n^2}\underbrace{\sum_{t=1}^{n}\cos^2(\omega_k t)}_{=\frac{n}{2}} = \tfrac{2\sigma^2}{n},$$

$$\mathrm{Var}(\hat{B}_k) = \ldots = \tfrac{4\sigma^2}{n^2}\underbrace{\sum_{t=1}^{n}\sin^2(\omega_k t)}_{=\frac{n}{2}} = \tfrac{2\sigma^2}{n},$$

$$\mathrm{Cov}(\hat{A}_j,\hat{B}_k) = \tfrac{4}{n^2}\sum_{s=1}^{n}\sum_{t=1}^{n}\mathrm{Cov}(x_s,x_t)\cos(\omega_j s)\sin(\omega_k t)$$

$$= \tfrac{4}{n^2}\sum_{t=1}^{n}Var(x_t)\cos(\omega_j t)\sin(\omega_k t)$$

$$= \frac{4\sigma^2}{n^2}\underbrace{\sum_{t=1}^{n}\cos(\omega_j t)\sin(\omega_k t)}_{=0} = 0.$$

Similarly, $\mathrm{Cov}(\hat{A}_j,\hat{A}_k) = \mathrm{Cov}(\hat{B}_j,\hat{B}_k) = 0$ if $j \neq k$. AM

Suppose that $x_1,\ldots,x_n$ are i.i.d. $N(\mu,\sigma^2)$. [1] Then

$$\hat{A}_k = \frac{2}{n}\sum_{t=1}^{n} x_t \cos\left(\frac{2\pi k}{n}t\right),\quad \hat{B}_k = \frac{2}{n}\sum_{t=1}^{n} x_t \sin\left(\frac{2\pi k}{n}t\right),$$

are for $1\le k\le m=[\frac{n-1}{2}]$ normally distributed with mean zero and variance $\frac{2\sigma^2}{n}$.

This implies that

$$\sqrt{\frac{n}{2\sigma^2}}\,\hat{A}_k\,,\quad \sqrt{\frac{n}{2\sigma^2}}\,\hat{B}_k$$

have a standard normal distribution. Because of the joint normality of $\hat{A}_k$, $\hat{B}_k$ it follows already from

$$\mathrm{Cov}(\hat{A}_k,\hat{B}_k)=0$$

that $\hat{A}_k$ and $\hat{B}_k$ are independent.

---

[1] When the sample size is large, this assumption is not critical. Because of the central limit theorem, the distributions of (weighted) averages of the observations can still be regarded as approximately normal even if the observations themselves are not normally distributed.

For each $1\le k\le m$, the random variable

$$\frac{n}{2\sigma^2}(\hat{A}_k^2+\hat{B}_k^2)=\frac{n}{2\sigma^2}\,\frac{8\pi}{n}\,\underbrace{\frac{n}{8\pi}(\hat{A}_k^2+\hat{B}_k^2)}_{I(\omega_k)}=\frac{4\pi}{\sigma^2}I(\omega_k)$$

is therefore the sum of the squares of two independent standard normal random variables, i.e., it has a chi-squared distribution with 2 degrees of freedom, denoted

$$\frac{4\pi}{\sigma^2}I(\omega_k)\sim\chi^2(2).$$

The random variables $I(\omega_k)$ are not only identically distributed, they are also independent. This follows from the independence of the pairs $(\hat{A}_k,\hat{B}_k)$, $k=1,\ldots,m$, and the fact that each $I(\omega_k)$ is a function of $\hat{A}_k$ and $\hat{B}_k$. AD

<u>Exercise:</u> Show that if $k=\frac{n}{2}$, then

$$\frac{2\pi}{\sigma^2}I(\omega_k)=\frac{2\pi}{\sigma^2}I(\pi)\sim\chi^2(1).$$ A1

If $x_1,\ldots,x_n$ are i.i.d. $N(\mu,\sigma^2)$, then

$$\frac{4\pi}{\sigma^2} I(\omega_k),\ 1\le k<\frac{n}{2},$$

are i.i.d. $\chi^2(2)$.

Since the $\chi^2(2)$ distribution is identical to the exponential distribution with mean 2,

$$\frac{2\pi}{\sigma^2} I(\omega_k),\ 1\le k<\frac{n}{2},$$

are i.i.d. Exp(1).

Hence,

$$P\left(\frac{2\pi}{\sigma^2} I(\omega_k)\le c\right)=\int_0^c e^{-\lambda}d\lambda=-e^{-\lambda}\Big|_0^c=1-e^{-c}=1-\alpha,$$

if $c=-\log(\alpha)$.

When n is large, we can approximate the unknown parameter σ^2 by the sample variance s^2.

The null hypothesis of Gaussian white noise can be rejected at the approximate 5% level, if for a specified k or for specified $k_1,\ldots,k_j$,

(i) $\frac{2\pi}{s^2} I(\omega_k) > -\log(0.05)$,

or (iii) $\frac{4\pi}{\sigma^2} I(\omega_{k_1})+\ldots+\frac{4\pi}{\sigma^2} I(\omega_{k_j}) > \chi^2_{0.95}(2j)$, AT

or (ii) $\max\left(\frac{2\pi}{s^2} I(\omega_{k_1}),\ldots,\frac{2\pi}{s^2} I(\omega_{k_j})\right) > -\log(1-\sqrt[j]{0.95})$,

where $\chi^2_{0.95}(2\cdot 1)\sim 5.991$, $\chi^2_{0.95}(2\cdot 2)\sim 9.488,\ldots$ are the 0.95-quantiles of the $\chi^2(2\cdot j)$ distribution, j=1,2,…

Exercise: Show that $Ex=2$ if $x\sim\chi^2(2)$. A2

Exercise: Suppose that $x_1,\ldots,x_n$ are i.i.d. $N(\mu,\sigma^2)$. Show that $E\, I(\omega_k)=\frac{\sigma^2}{2\pi}$ for all $1\le k<\frac{n}{2}$. AW

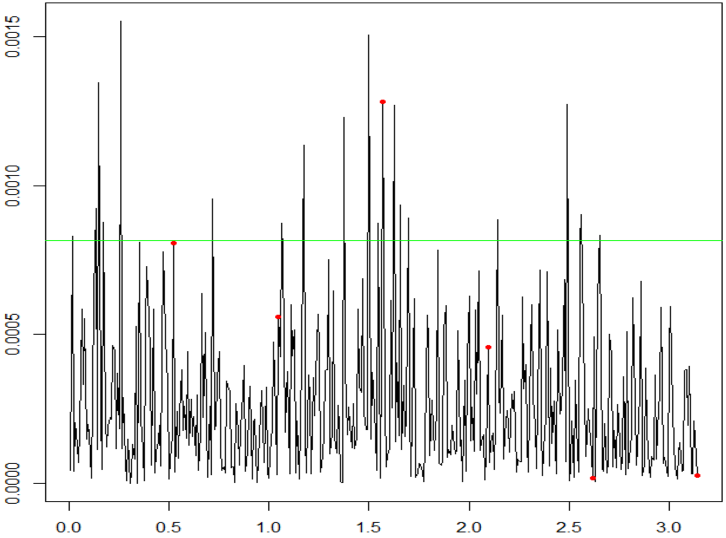
Exercise: Suppose that J_1, J_2 are i.i.d. Exp(1). Show that

$$P(\max(J_1,J_2)\le -\log(1-\sqrt{0.95}))=0.95.$$

Hint: $P(\max(J_1,J_2)\le c)=P(J_1\le c\wedge J_2\le c)=P(J_1\le c)P(J_2\le c)$ AL

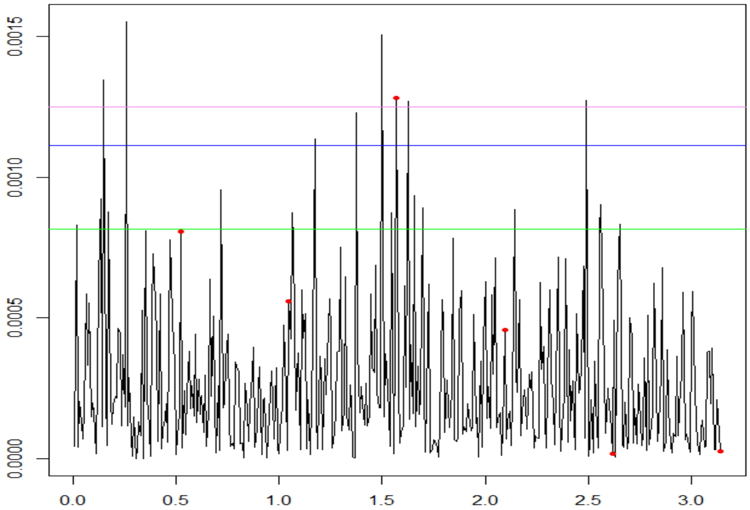
Exercise: Add a horizontal line to the S&P 500 periodogram, which represents the critical value (α=0.05) for a test based on the periodogram value at a fixed frequency.

```
s2 <- var(r12)
abline(h=-log(0.05)*s2/(2*pi),col="green")
```



Exercise: Add further horizontal lines to the S&P 500 periodogram, which represent the critical values (α=0.05) for tests based on the maximum of 3 or 5 periodogram values at 3 or 5 fixed frequencies.

```
abline(h=-log(1-0.95^(1/3))*s2/(2*pi),col="blue")
abline(h=-log(1-0.95^(1/5))*s2/(2*pi),col="violet")
```



Exercise: Use the sum of *j* periodogram values at fixed frequencies as test statistic.

(i) *j*=3: First three seasonal frequencies

```
sum(pg12[s[1:3]])*4*pi/s2 # test statistic
19.43515
qchisq(0.95,6) # 0.95-quantile of a chi2(6) distr.
12.59159
```

19.69437>12.59159 $\Rightarrow$ H_0 is rejected.

(ii) j=6: All seasonal frequencies

```
s <- n.years*(1:6)  # indices of seasonal frequencies
sum(pg12[s[1:5]])*4*pi/s2+pg12[s[6]]*2*pi/s2
23.01138
qchisq(0.95,11) # 0.95-quantile of a chi2(11) distr.
19.67514

h <- sum(pg12[s[1:5]])*4*pi/s2+pg12[s[6]]*2*pi/s2
1-pchisq(h,11)
0.0176099
```

Note: The periodogram at frequency π, which is the 6th seasonal frequency, has a different distribution under H_0, i.e.,

$$\frac{2\pi}{\sigma^2} I(\omega_k) \sim \chi^2(1) \text{ if } \omega_k = \pi$$

and

$$\frac{4\pi}{\sigma^2} I(\omega_k) \sim \chi^2(2) \text{ if } \omega_k \neq \pi.$$