Hochschule RheinMain | WS 2022/23
Prof. Dr. D. Lehmann

7. Übungsblatt zur Vorlesung
Datenanalyse mit R

Aufgabe 1) In den letzten beiden `week6.txt` und `week7.txt` hatten wir durch eine Analyse von Finanzzeitreihen das folgende Modell motiviert:

$$S(t_k) \;=\; S(t_{k-1}) \times \Big[\,1 \,+\, \mathrm{stddev}_d(t_{k-1})\,\phi_k\,\Big] \tag{1}$$

mit standard-normalverteilten Zufallszahlen $\{\phi_k\}_{k=1}^N$ und der d-Tages-Standardabweichung der Returns

$$\mathrm{stddev}_d^2(t_k) \;=\; \frac{1}{d}\Big\{\mathrm{ret}^2(t_k) + \mathrm{ret}^2(t_{k-1}) + \cdots + \mathrm{ret}^2(t_{k-d+1})\Big\} \tag{2}$$

Wir hatten dann einige Pfade simuliert. Wenn wir die $\mathrm{stddev}_d(t_{k-1})$ einfach durch eine Konstante ersetzt hatten, etwa $\mathrm{stddev}_d(t_{k-1}) = \mathrm{bsvol} = 1\%$ für alle t_k, dann sahen die Pfade ganz vernünftig aus. Wenn wir jedoch das $\mathrm{stddev}_d(t_k)$ mit der Formel (2) aus den bereits simulierten Daten berechnet hatten, und für das d etwa 20, 5 oder $d = 1$ gewählt hatten, dann wurden die Preispfade für die $S(t_k)$ sehr schnell konstant und die Volatilität ging sehr schnell nach 0. Das wollen wir hier jetzt etwas genauer analysieren. Dazu betrachten wir den Fall $d = 1$, wir bekommen dann das 'naive ARCH(1)-Modell' mit der Spezifikation

$$S(t_k) \;=\; S(t_{k-1}) \times \Big[\,1 \,+\, \mathrm{stddev}_1(t_{k-1})\,\phi_k\,\Big] \tag{3}$$

mit

$$\mathrm{stddev}_1^2(t_k) \;\overset{\substack{\text{Gleichung (2)}\\ \text{mit } d=1}}{=}\; \mathrm{ret}^2(t_k) \tag{4}$$

Die Returns $\mathrm{ret}(t_k)$ waren definiert durch

$$\mathrm{ret}(t_k) \;\overset{\text{allg. Def.}}{=}\; \frac{S(t_k) - S(t_{k-1})}{S(t_{k-1})} \;\overset{\text{hier, mit (3)}}{=}\; \mathrm{stddev}_1(t_{k-1})\,\phi_k \tag{5}$$

Wir können Gleichung (5) in Gleichung (4) einsetzen und erhalten

$$\begin{aligned}
\mathrm{stddev}_1^2(t_k) \;&=\; \mathrm{stddev}_1^2(t_{k-1})\,\phi_k^2 \\
&=\; \mathrm{stddev}_1^2(t_{k-2})\,\phi_{k-1}^2\,\phi_k^2 \\
&\;\;\vdots \\
&=\; \mathrm{stddev}_1^2(t_0)\,\phi_1^2\,\cdots\,\phi_{k-1}^2\,\phi_k^2 \\
&=:\; \mathrm{startvol}^2 \,\times\, \pi_k
\end{aligned} \tag{6}$$

wobei wir also die allererste $\mathrm{stddev}_1^2(t_0)$ gleich einer Konstanten setzen, die wir $\mathrm{startvol}^2$

genannt haben, etwa gleich $(1\%)^2$, und die Grössen π_k sind also gegeben durch

$$\pi_k \quad := \quad \phi_1^2\,\phi_2^2 \,\times\, \cdots \,\times\, \phi_{k-1}^2\,\phi_k^2 \quad . \tag{7}$$

Wir definieren ebenfalls noch die Grössen

$$s_k \quad := \quad \tfrac{1}{k}\left(\,\phi_1^2 \,+\, \phi_2^2 \,+\, \cdots \,+\, \phi_{k-1}^2 \,+\, \phi_k^2\,\right) \quad . \tag{8}$$

a) Zeigen Sie durch eine Rechnung mit Bleistift und Papier, ohne R-Code,

$$\begin{aligned} \mathsf{E}[\,s_k\,] \quad &= \quad 1 \\ \mathsf{E}[\,\pi_k\,] \quad &= \quad 1 \end{aligned}$$

für beliebiges k. Starten Sie dann eine R-Session und versuchen Sie, die folgenden Zahlen und Bilder zu reproduzieren.

b) Wir wollen jeweils

$$N \quad = \quad 10000$$

Realisationen der Zufallszahlen

$$\begin{aligned} s_5\;,\;s_{10}\;,\;s_{15}\;,\;s_{20} \quad &\in \quad \mathbb{R} \\ \pi_5\;,\;\pi_{10}\;,\;\pi_{15}\;,\;\pi_{20} \quad &\in \quad \mathbb{R} \end{aligned}$$

generieren und in den Vektoren

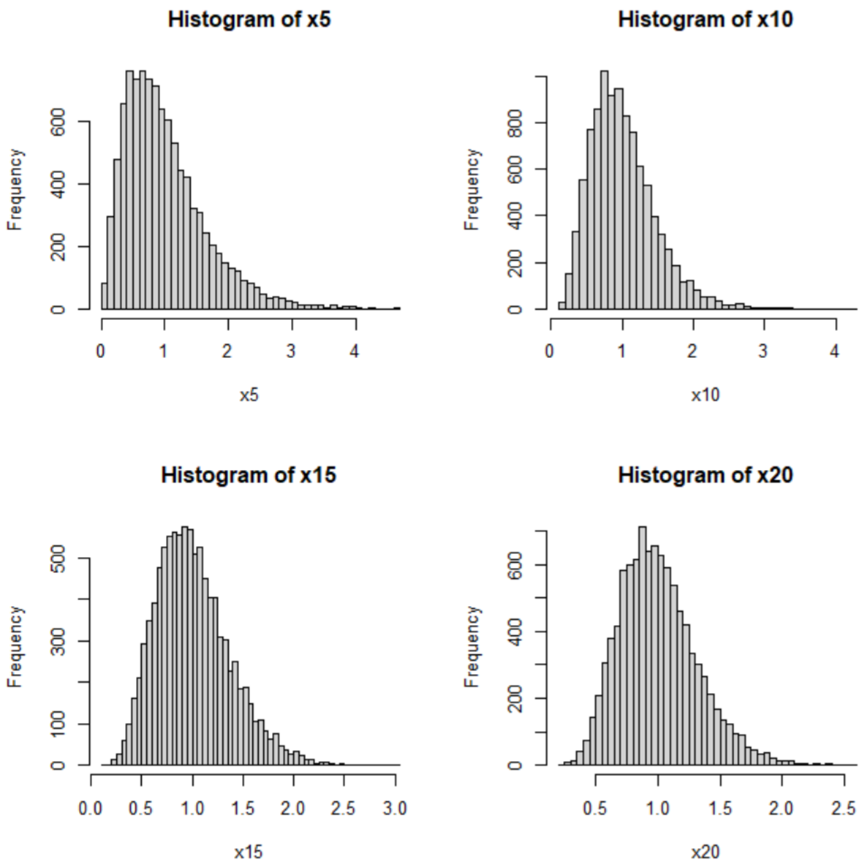
$$\begin{aligned} x_5\;,\;x_{10}\;,\;x_{15}\;,\;x_{20} \quad &\in \quad \mathbb{R}^N \\ y_5\;,\;y_{10}\;,\;y_{15}\;,\;y_{20} \quad &\in \quad \mathbb{R}^N \end{aligned}$$

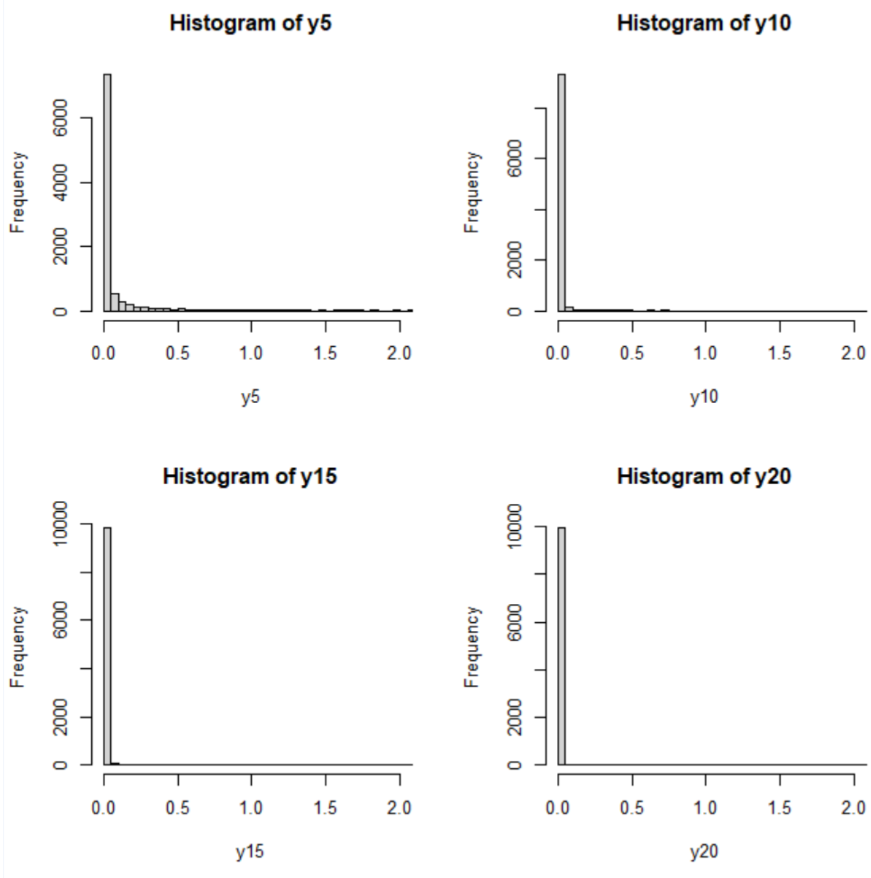
abspeichern. Schauen Sie sich dann zunächst mal die Quartile, Mediane und Mittelwerte dieser jeweils 10000 Zahlen an:

```
> summary(x5)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
0.01088 0.53266 0.86840 0.99892 1.31642 4.68862
> summary(x10)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.1024  0.6781  0.9374  1.0023  1.2507  4.2755
> summary(x15)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.1434  0.7389  0.9628  1.0052  1.2217  3.0469
> summary(x20)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.2296  0.7784  0.9703  1.0029  1.1876  2.5541
```

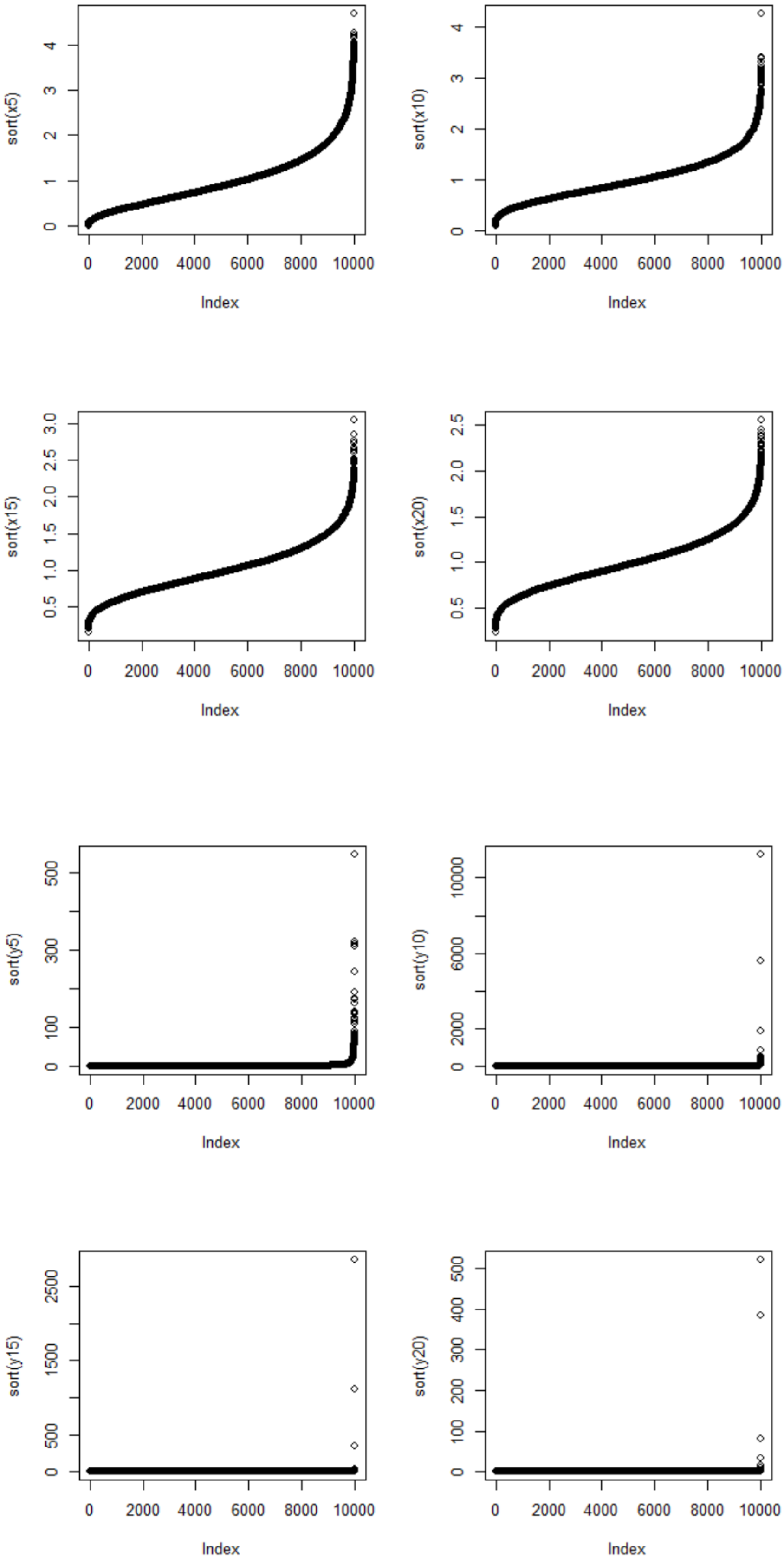
```
> summary(y5)
    Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
  0.0000   0.0001   0.0029   1.0133   0.0616 548.4202
> summary(y10)
     Min.   1st Qu.    Median      Mean   3rd Qu.      Max.
    0.000     0.000     0.000     2.421     0.000 11265.759
> summary(y15)
     Min.   1st Qu.    Median      Mean   3rd Qu.      Max.
   0.0000    0.0000    0.0000    0.5754    0.0000 2854.4720
> summary(y20)
    Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
  0.0000   0.0000   0.0000   0.1089   0.0000 520.7419
```

c) Schauen wir uns jetzt die Histogramme, die Verteilungen dieser jeweils 10000 Zufallszahlen an:





d) Ordnen Sie dann diese jeweils 10000 Zufallszahlen der Grösse nach und plotten Sie die geordneten Zahlen. Man bekommt etwa das folgende:



e) Schliesslich wollen wir uns noch die Monte Carlo Summen für die Erwartungswerte der s_k und der π_k anschauen. Für die s_k bekommt man in wesentlichen immer dasselbe Bild, aber die Bilder für die π_k können recht variable sein, je nachdem, wie die Zufallszahlen sich so realisiert haben: