

linux: arkistointi

- ▶ **tar** liittää useampia tiedostoja yhteen samaan arkistoon (engl. archive)

Esimerkki 1

tar cvf arkisto.tar *.DAT

luo arkiston **arkisto.tar**, joka sisältää kaikki **.DAT** loppuiset tiedostot

- ▶ **tar -xvf arkisto.tar**
⇒ arkisto takaisin yksittäisiksi tiedostoiksi

Esimerkki 2

tar cvzf arkisto.tar.gz *.DAT

luo kompressoidun (engl. compress) arkiston, joka vie vähemmän levytilaa

- ▶ **tar -xvzf arkisto.tar.gz**
purkaa arkiston yksittäisiksi tiedostoiksi
Useimmille nämä riittävät. Lisää esimerkkejä löytyy vaikkapa [täältä](#) | [www](#) |
- ▶ **Kuva:** **arkisto.tar** luodaan **/nyt** hakemistoon, kopioidaan **/nyt1** hakemistoon, missä se puretaan
- ▶ **Esimerkkikäyttö:** Lähettäjä luo ja lähettää arkiston, jonka vastaanottaja purkaa

```
jetsu@dx5-flspa-02: ~/nyt1
File Edit View Search Terminal Help
jetsu@dx5-flspa-02:~/nyt1$ ls
jetsu@dx5-flspa-02:~/nyt1$ touch aaa.dat
jetsu@dx5-flspa-02:~/nyt1$ touch bbb.dat
jetsu@dx5-flspa-02:~/nyt1$ touch ccc.dat
jetsu@dx5-flspa-02:~/nyt1$ ls
aaa.dat bbb.dat ccc.dat
jetsu@dx5-flspa-02:~/nyt1$ tar cvf arkisto.tar *.dat
aaa.dat
bbb.dat
ccc.dat
jetsu@dx5-flspa-02:~/nyt1$ cd ..
jetsu@dx5-flspa-02:~$ mkdir nyt1
jetsu@dx5-flspa-02:~$ cd nyt1
jetsu@dx5-flspa-02:~/nyt1$ ls
jetsu@dx5-flspa-02:~/nyt1$ cd ..
jetsu@dx5-flspa-02:~$ cd nyt
jetsu@dx5-flspa-02:~/nyt$ cp arkisto.tar /home/jetsu/nyt1/
jetsu@dx5-flspa-02:~/nyt$ cd ..
jetsu@dx5-flspa-02:~$ cd nyt1
jetsu@dx5-flspa-02:~/nyt1$ ls
arkisto.tar
jetsu@dx5-flspa-02:~/nyt1$ tar -xvf arkisto.tar
aaa.dat
bbb.dat
ccc.dat
jetsu@dx5-flspa-02:~/nyt1$ ls
aaa.dat arkisto.tar bbb.dat ccc.dat
jetsu@dx5-flspa-02:~/nyt1$
```

linux: koneelta toiselle

- ▶ **scp** komento
- ▶ **Esimerkki 1:** Työskentelet koneessa **sky1** hakemistossa **~username/**. Komento

```
scp *.DAT sky2:/home/username/dir1/
kopioi kaikki työhakemistosi .DAT
päätteiset tiedostot toisen koneen sky2
hakemistoon /home/username/dir1/
Huom: sky2 kysyy salasanaa!
Selvennys: sky1 tiedostot → sky2
```

linux: koneelta toiselle

- ▶ **Esimerkki 2:** Työskentelet koneessa **sky1** hakemistossa **~user/dir2/**. Komento
- ```
scp user@sky2:~user/dir1/*.PRO ~user/dir2/
kopioi kaikki koneen sky2 hakemistossa
~user/dir1/ olevat .PRO loppuiset tiedostot
koneen sky1 hakemistoon ~user/dir2/
```
- Huom:** sky2 kysyy salasanaa!
- Selvennys:** username lyhennetty user
- Selvennys:** sky2 tiedostot → sky1
- ▶ Jälleen loputtomasti **vaihtoehtoja** | [www](http://www) |
  - ▶ Paljon dataa: Ensin **tar**. Sitten **scp**

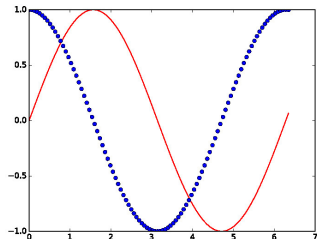
```
jetsu@dx5-flspa-02: ~/nyt1
File Edit View Search Terminal Help

jetsu@dx5-flspa-02:~/nyt1$ ls
aaa.dat arkisto.tar bbb.dat ccc.dat
jetsu@dx5-flspa-02:~/nyt1$ scp *.dat cedi.it.helsinki.fi:~jetsu/public_html/
jetsu@cedi.it.helsinki.fi's password:
aaa.dat 100% 0 0.0KB/s 00:00
bbb.dat 100% 0 0.0KB/s 00:00
ccc.dat 100% 0 0.0KB/s 00:00
jetsu@dx5-flspa-02:~/nyt1$
```

# Datan graafinen esitys

```
#
Kommenttirivi: Tama on python ohjelmani Plottimalli1.py
– Plottaa ensimmäisen kerran kuvan
#
import os ; os.system('clear') # Tyhjennetaan näyttö
import numpy as np # numpy importoitu
import pylab as pl # pylab importoitu
x = np.pi*np.arange(100)/49.0 # x-vektori
y = np.sin(x) # y-vektori
z = np.cos(x) # z-vektori
pl.plot(x,y, 'r') # x,y plotti
pl.plot(x,z, 'ob') # x,z plotti
pl.savefig('Plotti1.eps') # Tallenna eps-tiedostoon
```

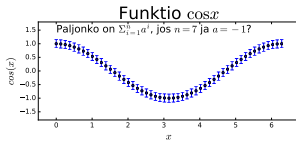
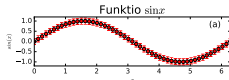
- ▶ **import pylab** mahdollistaa plottaamisen
  - ▶ **pl.plot** komento plottaa
  - ▶ **'r'** = punainen jatkuva viiva
  - ▶ **'ob'** = ympyrä **o** on sininen **b**
  - ▶ **pl.savefig...** tallentaa oikealla näkyvään plotin
- ⇒ Katsotaan plottia **evince Plotti1.eps**



# Datan graafinen esitys

```
Kommenttirivi: Tama on ohjelmani Plottimalli2.py
#
import os ; os.system('clear') # Tyhjennetaan naytto
import numpy as np # Importoi numpy
import pylab as pl # --"--- pylab
x = np.pi*np.arange(51)/25.0 # x-vektori
y = np.sin(x) # y-vektori
z = np.cos(x) # z-vektori
e = np.zeros(len(y))+0.15 # virhe
#
pl.axes([0.3,0.7,0.6,0.2]) # 1. plotin tila: pienempi
pl.xlim([0.0, 2.0 * np.pi]) # x rajat (vahan tilaa)
pl.ylim([-1.2, 1.2]) # y rajat (vahan tilaa)
pl.errorbar(x,y,e,fmt='or') # plottaa x,y,e
pl.title('Funktio_{$\sin(x)$}',fontsize=20,) # Paateksti
pl.xlabel('x',fontsize=10) # x-teksti
pl.ylabel('$\sin(x)$',fontsize=10) # y-teksti
pl.text(5.6,0.7,'(a)',fontsize=15) # Kirjoitetaan jotain
#
pl.axes([0.15,0.1,0.80,0.4]) # 2. plotin tila: suurempi
pl.xlim([-0.5, 2.0 * np.pi+0.5]) # x rajat (paljon tilaa)
pl.ylim([-1.8, 1.8]) # y rajat (paljon tilaa)
pl.errorbar(x,z,e,fmt='ob',ms=4) # plottaa x,z,e
pl.title('Funktio_{$\cos(x)$}',fontsize=30) # Paateksti
pl.xlabel('x',fontsize=18) # x-teksti
pl.ylabel('$\cos(x)$',fontsize=18) # y-teksti
txt='Paljonko on $\Sigma_{i=1}^n a^i$, jos $n=7$ ja $a=-1$?'
pl.text(0.0,1.4,txt,fontsize=18) # Kirjoitetaan jotain
#
pl.savefig('Plotti2.pdf') # Kuva tiedosto
```

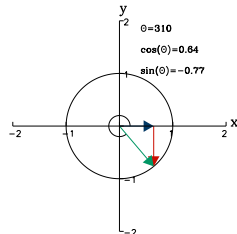
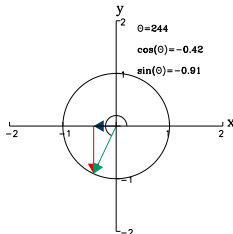
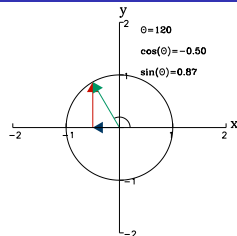
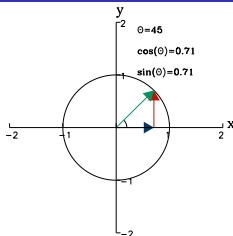
- Kotisivulta **Plottimalli2.py**
- **Plotti2.pdf**: Eri kokoiset kuvat  
**pl.axes** kuvien paikka ja koko  
**pl.xlim** x-rajat, **pl.ylim** y-rajat  
**pl.errorbar** x,y,e (e=error bars)  
**pl.title** pääteksti (**L<sup>A</sup>T<sub>E</sub>X**)  
**pl.xlabel** x-teksti (**L<sup>A</sup>T<sub>E</sub>X**)  
**pl.ylabel** y-teksti (**L<sup>A</sup>T<sub>E</sub>X**)  
**pl.text** asemoitu teksti (**L<sup>A</sup>T<sub>E</sub>X**)



## Rayleigh testi

## Yksikkövektorit

- ▶ Yksikkövektorin pituus =  $|\vec{r}| = 1$
  - ▶ Yksikkövektorin suunta = Vaihekulma =  $\Theta$
  - ▶ Yksikkövektorit alkavat origosta
  - ▶ Yksikkövektorit osoittavat yksikköympyrälle
  - ▶  $\vec{r} = [|\vec{r}| \cos \Theta, |\vec{r}| \sin \Theta] = [\cos \Theta, \sin \Theta]$
  - ▶  $\vec{x} = [\cos \Theta, 0]$
  - ▶  $\vec{y} = [0, \sin \Theta]$
  - ▶  $\vec{r} = \vec{x} + \vec{y}$
  - ▶  $\cos \Theta = x/|\vec{r}| \Leftrightarrow x = \cos \Theta$
  - ▶  $\sin \Theta = y/|\vec{r}| \Leftrightarrow y = \sin \Theta$
  - ▶ Pythagoras:  $|\vec{r}|^2 = |\vec{x}|^2 + |\vec{y}|^2$
- $$\Leftrightarrow 1^2 = (\cos \Theta)^2 + 0^2 + 0^2 + (\sin \Theta)^2$$



# Rayleigh testi

$\bar{R}$  = Summa  $n$ :stä yksikkövektorista  $\bar{r}_i$

- ▶  $\bar{R} = \sum_{i=1}^{i=n} \bar{r}_i = \bar{r}_1 + \bar{r}_2 + \dots + \bar{r}_n$
- ▶  $\bar{r}_1 = \bar{x}_1 + \bar{y}_1, \bar{r}_2 = \bar{x}_2 + \bar{y}_2, \dots, \bar{r}_n = \bar{x}_n + \bar{y}_n$   
 $\Rightarrow \bar{R} = \bar{x}_1 + \bar{x}_2 + \dots \bar{x}_n + \bar{y}_1 + \bar{y}_2 + \dots \bar{y}_n$
- ▶  $\bar{R} = \bar{X} + \bar{Y}$
- ▶  $\bar{X} = \sum_{i=1}^{i=n} \bar{x}_i = \sum_{i=1}^{i=n} [\cos \Theta_i, 0] = [\cos \Theta_1 + \cos \Theta_2 + \dots + \cos \Theta_n, 0] = [\sum_{i=1}^{i=n} \cos \Theta_i, 0]$
- ▶  $\bar{Y} = \sum_{i=1}^{i=n} \bar{y}_i = \sum_{i=1}^{i=n} [0, \sin \Theta_i] = [0, \sin \Theta_1 + \sin \Theta_2 + \dots + \sin \Theta_n] = [0, \sum_{i=1}^{i=n} \sin \Theta_i]$
- ▶  $|\bar{X}|^2 = (\sum_{i=1}^{i=n} \cos \Theta_i)^2 + 0^2 = (\sum_{i=1}^{i=n} \cos \Theta_i)^2$
- ▶  $|\bar{Y}|^2 = 0^2 + (\sum_{i=1}^{i=n} \sin \Theta_i)^2 = (\sum_{i=1}^{i=n} \sin \Theta_i)^2$
- ▶ Pythagoras  $\Rightarrow |\bar{R}|^2 = |\bar{X}|^2 + |\bar{Y}|^2$
- ▶ Lopputuloks  $|\bar{R}|^2 = (\sum_{i=1}^{i=n} \cos \Theta_i)^2 + (\sum_{i=1}^{i=n} \sin \Theta_i)^2$
- ▶ **Rayleigh:n testiparametri**  $n$ :lle vaihekulmalle  $\Theta_1, \Theta_2, \dots, \Theta_n$  on

$$z = \frac{|\bar{R}|^2}{n} = \frac{(\sum_{i=1}^{i=n} \cos \Theta_i)^2 + (\sum_{i=1}^{i=n} \sin \Theta_i)^2}{n}$$

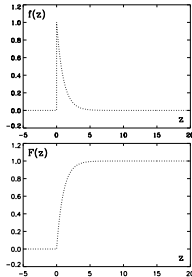
- $$f(z) = \begin{cases} 0, & z < 0 \\ e^{-z} & z \geq 0 \end{cases}$$

- $$F(z_0) = P(z \leq z_0) = \int_{-\infty}^{z_0} f(x) dz = \int_{-\infty}^0 0 dz + \int_0^{z_0} e^{-z} dz \\ = 0 + \int_0^{z_0} e^{-z} dz = -e^{-z} - (-e^0) = 1 - e^{-z_0}$$

$$F(z) = \begin{cases} 0, & z < 0 \\ 1 - e^{-z}, & z \geq 0, \end{cases}$$

- ▶  $P(z \leq z_0)$  on todennäköisyys, että  $z$  on valittua  $z_0$  arvoa pienempi
- ▶ Komplementti tapaus:  

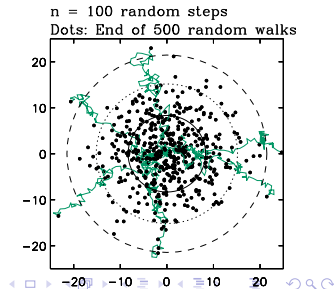
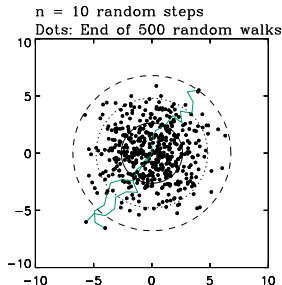
$$P(z > z_0) = 1 - P(z \leq z_0) = 1 - (1 - e^{-z_0}) = e^{-z_0}$$
- ▶ **Esimerkki:**  $0.5 = P(z \leq z_0) = 1 - e^{-z_0} \Leftrightarrow e^{-z_0} = 0.5 \Leftrightarrow -z_0 = \ln 2^{-1} \Leftrightarrow z_0 = \ln 2 \approx 0.693$ . Tarkoittaa, että puolet  $z$  arvoista välillä  $0 \leq z < 0.693$ , eli toinen puoli on välillä  $0.693 < z < n$ .



# Satunnaiskulku

- ▶ Aloita satunnaiskulku origosta: **Ota  $n$  askelta**  $\bar{r}_1, \bar{r}_2, \dots, \bar{r}_n$
- ▶ Jokaisen askeleen pituus on  $|r_i| = \sqrt{\cos^2 \Theta_i + \sin^2 \Theta_i} = 1$
- ▶ Valitse jokaisen askeleen suunta  $\Theta_i$  satunnaisesti
- ▶ **Ongelma:** Kuinka kauas origosta todennäköisesti pääset?
- ▶  $Q = P(z > z_0) = 1 - P(z \leq z_0) = 1 - (1 - e^{-z_0}) = e^{-z_0} \Leftrightarrow Q = e^{-z_0} \Leftrightarrow \ln Q = -z_0 \Leftrightarrow \mathbf{z_0 = -\ln Q}$
- ▶ Toisaalta  $\mathbf{z_0 = \frac{|\bar{R}_0|^2}{n}} \Rightarrow |\bar{R}_0| = \sqrt{n z_0} = \sqrt{-n \ln Q}$
- ▶ **Kuva:** Etäisyydet  $|\bar{R}_0|$  ratkaistu  $n = 10$  ja  $n = 100$  askeleelle
- ▶ **Todennäköisyydet** ovat  $Q = 0.5$  (puolet tapauksista),  $Q = 0.1$  (yhden kerran kymmenestä) ja  $Q = 0.01$  (yhden kerran sadasta)
- ▶ Pisteet ovat 500 satunnaiskulun päätepisteitä. Tapausten  $Q \leq 0.01$  reitit näytetty **vihreän värinä**

|           | $Q = 0.5$                            | $Q = 0.1$                             | $Q = 0.01$                            |
|-----------|--------------------------------------|---------------------------------------|---------------------------------------|
| $n = 10$  | $z_0 = 0.69$<br>$ \bar{R}_0  = 2.63$ | $z_0 = 2.30$<br>$ \bar{R}_0  = 4.80$  | $z_0 = 4.60$<br>$ \bar{R}_0  = 6.79$  |
| $n = 100$ | $z_0 = 0.69$<br>$ \bar{R}_0  = 8.33$ | $z_0 = 2.30$<br>$ \bar{R}_0  = 15.17$ | $z_0 = 4.60$<br>$ \bar{R}_0  = 21.46$ |
|           | jatkuva                              | pisteitä                              | tavuviivoja                           |





- ▶ **Rayleigh testin idea:** “Oletetaan, että aika on pitkä suora viiva. Tapahtumia edustavat pisteet, jotka ovat viivalla tasavälisillä etäisyyksillä  $2\pi$ . Kelaataan aika pyörälle, jonka läpimitta on  $d = 1$ . Tapahtumia edustavat pisteet päätyvät samaan paikkaan pyörän päälle. Kelataan aika takaisin suoraksi viivaksi ja poistetaan osa näistä tapahtumia edustavista pisteistä, jolloin tapahtumien väliin syntyy aukkoja. Jos aika kelataan taas takaisin pyörälle, pisteet päätyvät edelleen samaan paikkaan pyörän päälle. Mutta jos pyörän läpimitta on  $d \neq 1$ , pisteet eivät päädy samaan paikkaan, vaan ne leviävät satunnaisesti ympäri pyörän pintaa. Tämä on analogia Rayleigh:n testistä. Testissä projisoidaan aikapisteet yksikköympyrälle käyttäen testattavan periodin arvoa  $P$ . Aikapisteet päätyvät samaan suuntaan yksikköympyrälle, jos niiden jakauma on säännöllinen testattavalla periodilla  $P$ .”

- ▶ Ajan hetket  $t_i$  ( $i = 1, 2, 3, \dots, n$ )
- ▶  $\Theta_i = 2\pi t_i$  on vaihekulma
- ▶ Yksikkövektorien  $\vec{r}_i = [\cos \Theta_i, \sin \Theta_i]$  summa on  $\vec{R} = \sum_{i=1}^{i=n} \vec{r}_i$
- ▶ Summavektorin pituuden neliö on  $|\vec{R}|^2 = (\sum_{i=1}^{i=n} \cos \Theta_i)^2 + (\sum_{i=1}^{i=n} \sin \Theta_i)^2$
- ▶ **Rayleigh:n testiparametri** on

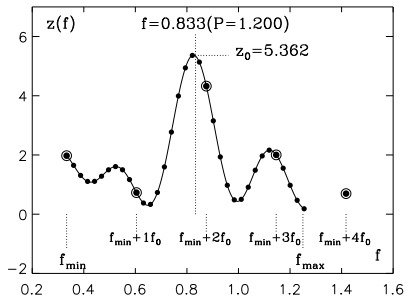
$$z = \frac{|\vec{R}|^2}{n} = \frac{(\sum_{i=1}^{i=n} \cos \Theta_i)^2 + (\sum_{i=1}^{i=n} \sin \Theta_i)^2}{n}$$

- ▶  $\vec{r}_i$  osoittavat saamaan suuntaan  $\Rightarrow |\vec{R}|$  suuri  $\Rightarrow z$  suuri  $\Rightarrow$  **Periodisuutta**
- ▶  $\vec{r}_i$  osoittavat satunnaisesti suuntiin  $\Rightarrow |\vec{R}|$  pieni  $\Rightarrow z$  pieni  $\Rightarrow$  **Ei periodisuutta**

## Paras periodi?

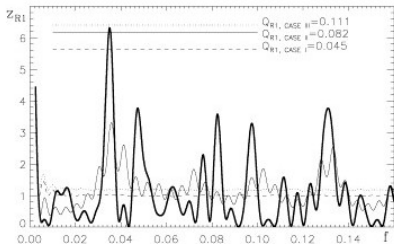
- ▶ Testataan monta periodia
- ▶ **Kuva:** testattu frekvenssien väli  $f_{\min} = 1/P_{\max} = 0.33$ ,  $f_{\max} = 1/P_{\min} = 1.25$  vastaa periodi väliä  $P_{\min} = 0.8$  ja  $P_{\max} = 3.0$
- ▶  $f_0 = 1/\Delta T$  riippumattomien frekvenssien  $f$  välinen etäisyys missä  $\Delta T = t_n - t_1$  (mustat  $\odot$  symbolit)
- ▶ “Testattavien frekvenssien ylitäyttö” muuttujalla  $\text{OFAC}=10 \Rightarrow$  Tiheämmät testattavien frekvenssien arvot  $f_{\text{step}} = f_0/\text{OFAC}$  (mustat “•” symbolit)
- ▶ Vielä tiheämpi testattujen frekvenssien väli = jatkuva viiva
- ▶ Maksimi arvo  $z(f) = z_0 = 5.362 =$  korkein periodogrammin huippu **parhaalla periodilla**  
 $P = 1.2 = f = 0.833$

- ▶  $Q = P(z > z_0) = e^{-z_0}$  on todennäköisyys, että  $z$  **ylittää** valitun arvon  $z_0$  **yhdessä  $m = 1$  testissä**
- ▶  $1 - Q$  on todennäköisyys, että  $z$  **ei ylitä** valittua arvoa  $z_0$  **yhdessä  $m = 1$  testissä**
- ▶  $(1 - Q)^m$  on todennäköisyys, että  $z$  **ei ylitä kertaakaan** valittua arvoa  $z_0$   **$m > 1$  testissä**
- ▶  $Q' = 1 - (1 - Q)^m$  on todennäköisyys, että  $z$  **ylittää** valitun arvon  $z_0$  **ainakin kerran  $m > 1$  testissä**
- ▶ Kuvassa noin neljä riippumatonta testiä =  $\odot$  välejä



# Rayleigh test

- ▶ “Spurious periods in the terrestrial impact crater record” (Jetsu, L., Pelt, J., 2000, A&A 353, 409) | [www.ijl.fi](http://www.ijl.fi)
- ▶ **Main result:** No periodicity in terrestrial impact cratering rate – Previous detections spurious



# Rayleigh test

- ▶ “Did the Ancient Egyptians Record the Period of the Eclipsing Binary Algol—The Raging One?” (Jetsu et al. 2013, *ApJ* 773, 1) | [www](#) |
- ▶ “Shifting Milestones of Natural Sciences: The Ancient Egyptian Discovery of Algol's Period Confirmed” (Jetsu et al. 2015, *PLOS ONE*, 10 (12) e0144140) | [www](#) |
- ▶ “Algol as Horus in the Cairo Calendar: The Possible Means and the Motives of the Observations” (Porceddu et al. 2018, *Open Astronomy* 27, 232) | [www](#) |

- ▶ **Main result:** Ancient Egyptians discovered the first variable star and determined its period 3000 years before modern astronomers. Algol was called Horus.

