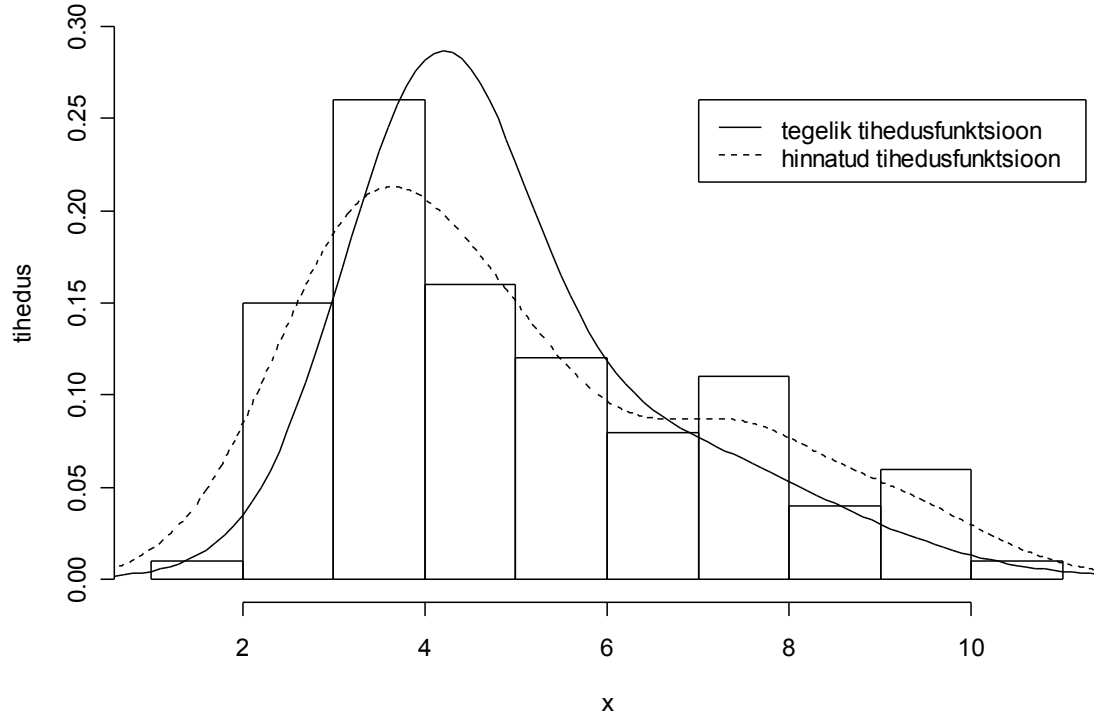


## Biomeetria bioloogidele 3. praktikum

**Näide: Tegelik jaotus, hinnatud jaotus ja jaotust kirjeldavad statistikumid**

**Valimi (n=100) histogramm, hinnatud tihedusfunktsioon ja populatsiooni tegelik tihedusfunktsioon**



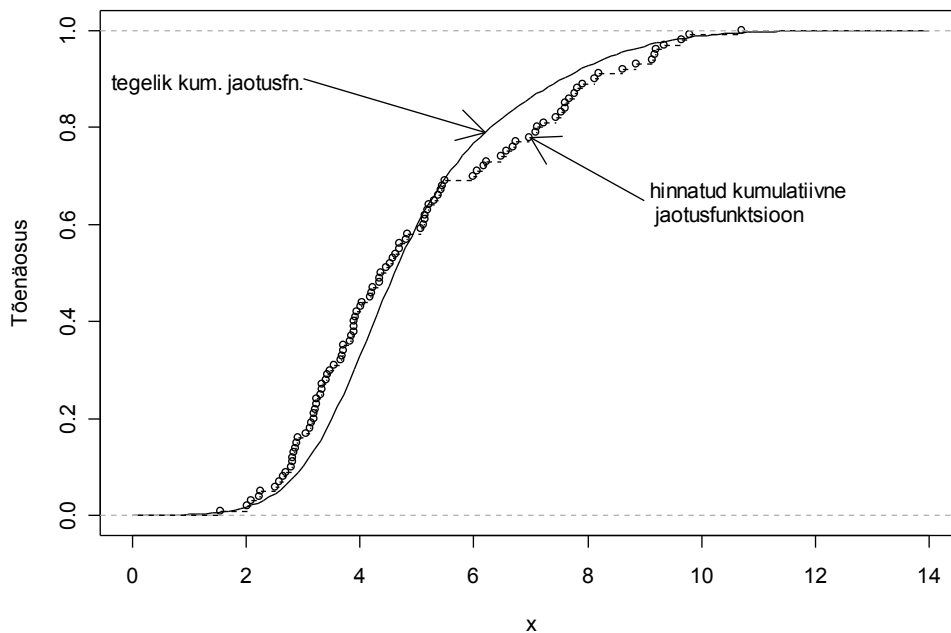
Tegelik (populatsiooni):

keskväärtus  $EX = 4,94$   
 mediaan  $med = 4,60\dots$   
 standardhälve  $\sigma = 1,80\dots$

Valimi:

keskmine  $\bar{x} = 5,02\dots$   
 mediaan  $med = 4,40\dots$   
 standardhälve  $s = 2,14$

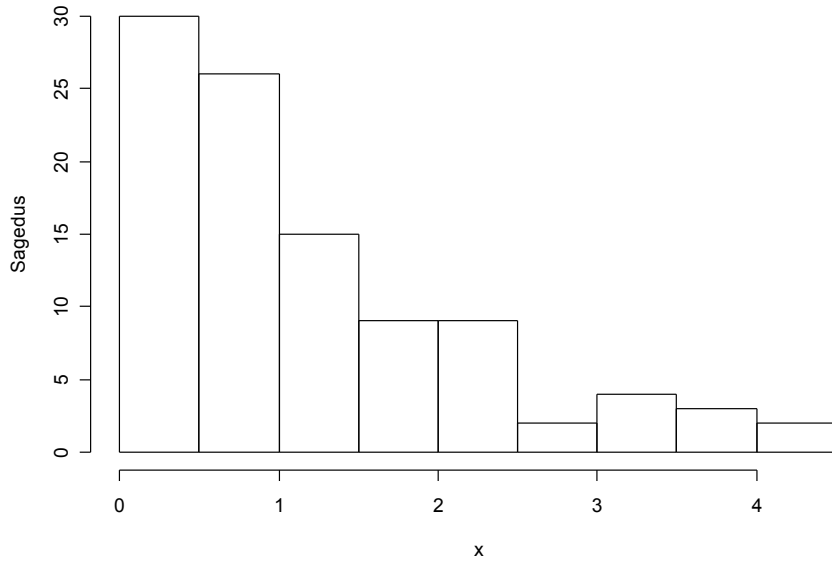
**kumulatiivne jaotusfunktsioon,  
tegelik ja hinnatud**



# Ülesanne 1

Kasutades jooniseid ürita ära arvata statistikute väärtuseid

1. pilt (histogramm)

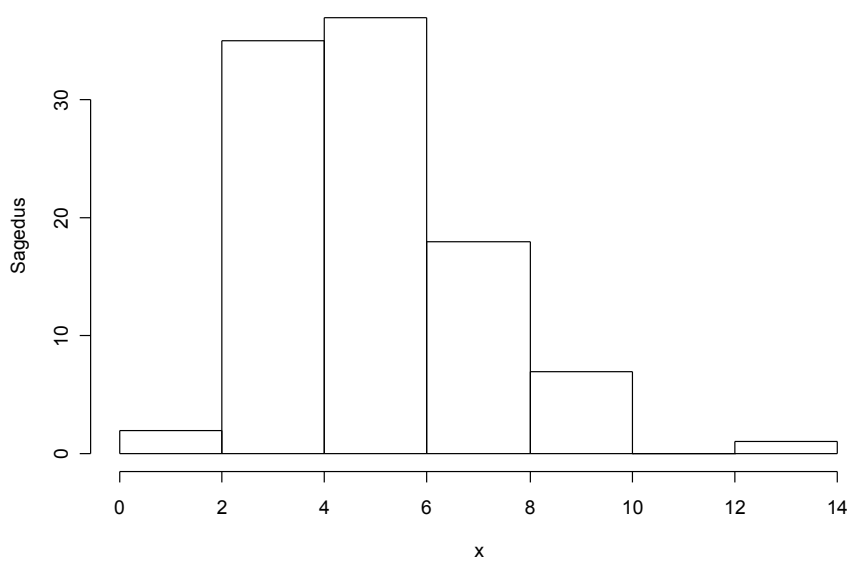


$\bar{x}$  = .....

med = .....

s = .....

2. pilt (histogramm)

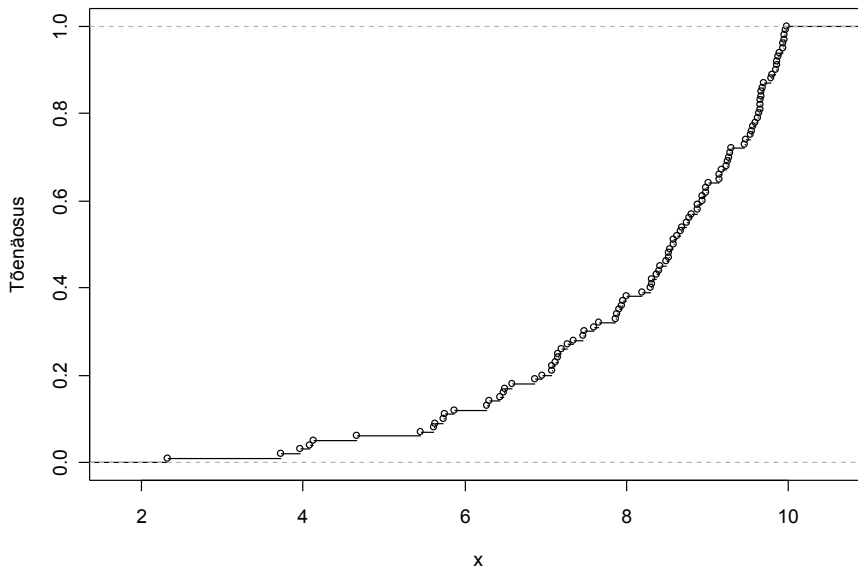


$\bar{x}$  = .....

med = .....

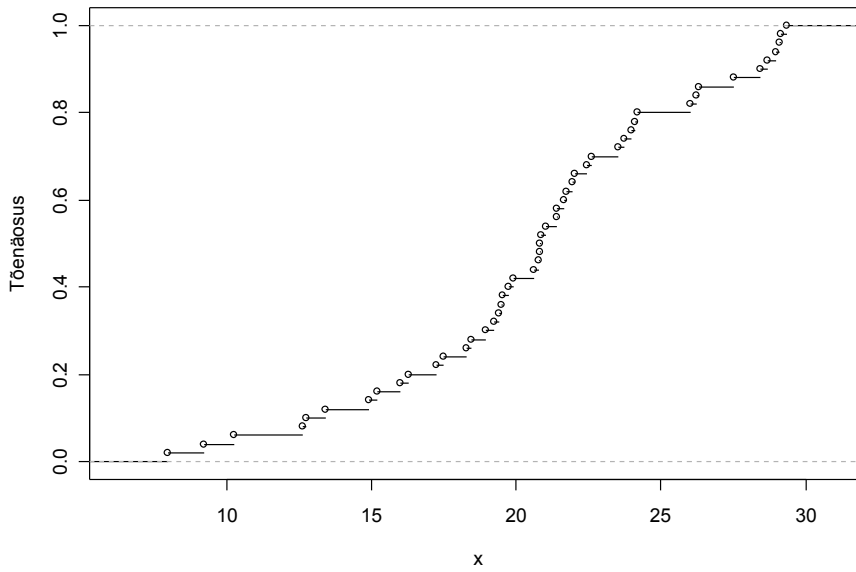
s = .....

3. pilt (kumulatiivne jaotusfunktsioon)



$\bar{x}$  = .....  
 med = .....  
 s = .....  
 0,05-kvantiil = .....  
 0,95-kvantiil = .....

4. pilt (kumulatiivne jaotusfunktsioon)

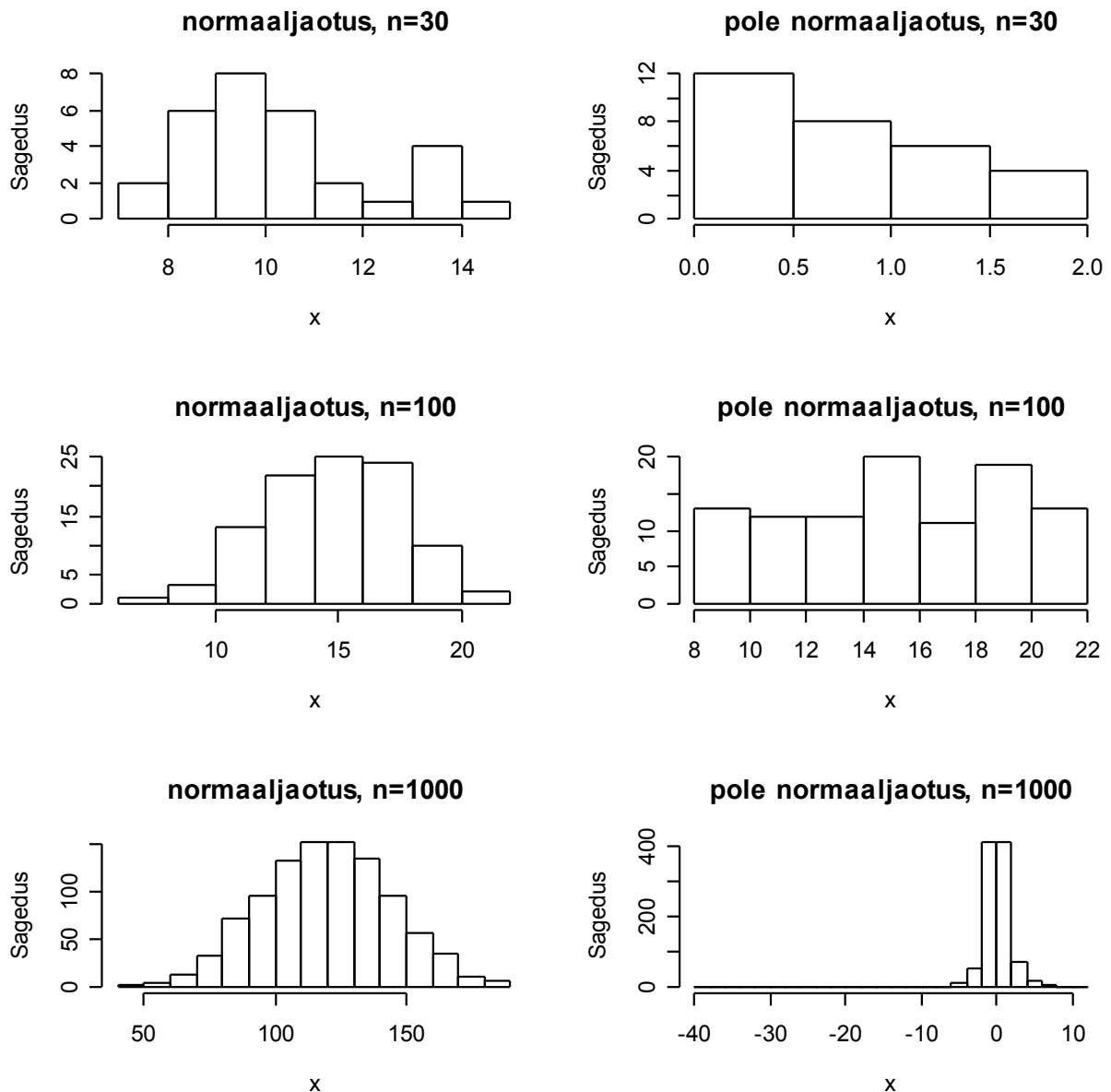


$\bar{x}$  = .....  
 med = .....  
 s = .....  
 0,05-kvantiil = .....  
 0,95-kvantiil = .....

## Kas uuritav tunnus on normaaljaotusega?

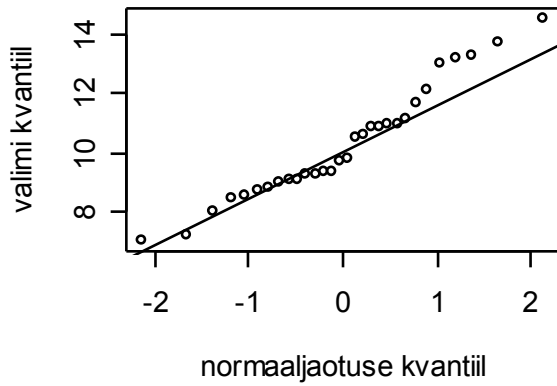
Vahel tuleb otsustada, kas uuritav tunnus võiks olla normaaljaotusega või mitte. Eriti väiksema valimi korral võib taolisele küsimusele vastamine keerukaks osutuda. Vaatame mõningaid võimalusi sellele küsimusele vastamiseks.

Üheks võimaluseks oleks teha otsus uuritava tunnuse histogrammi põhjal. Näiteid:

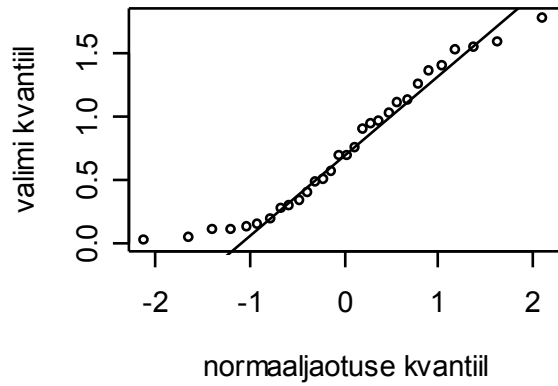


Histogramm ei pruugi olla kõige mugavam vahend normaaljaotuse tuvastamiseks. Kasutatakse ka tõenäosuspaberit (quantile-quantile plot, qq-plot) kus moodustatakse valimi kvantiilidest ja normaaljaotuse vastavatest kvantiilidest punktipaarid. Punktid kantakse graafikule ja juhul, kui uuritav tunnus on ka tegelikult normaaljaotusega, peaksid kõik punktid jääma enam-vähem ühele sirgele. Kui punktid kipuvad märgatavalt sirgest kõrvale kalduma, siis pole kardetavasti tegemist normaaljaotusega. Järgneval joonisel on antud tõenäosuspaberid kõigi 6 juhu tarvis, mille jaoks joonistasime ka histogrammid.

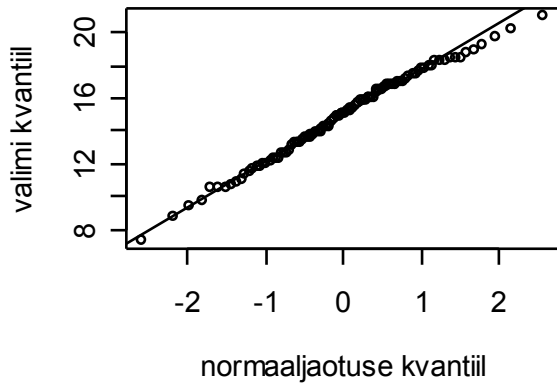
**normaaljaotus, n=30**



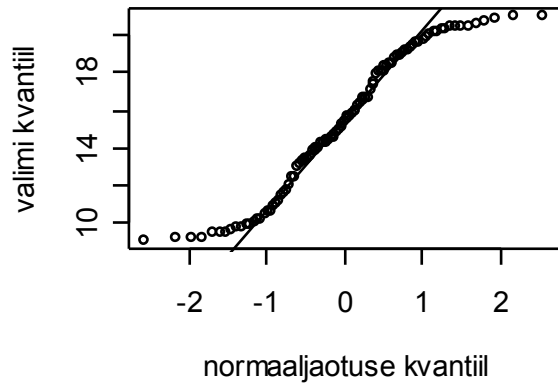
**pole normaaljaotus, n=30**



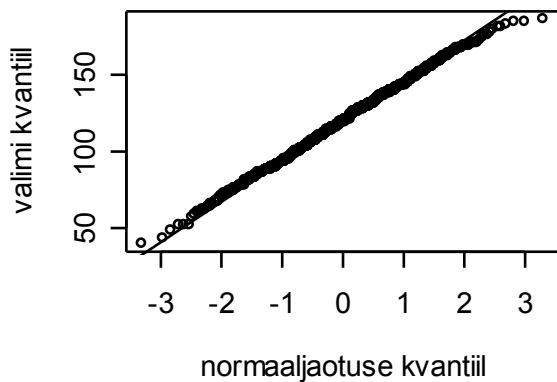
**normaaljaotus, n=100**



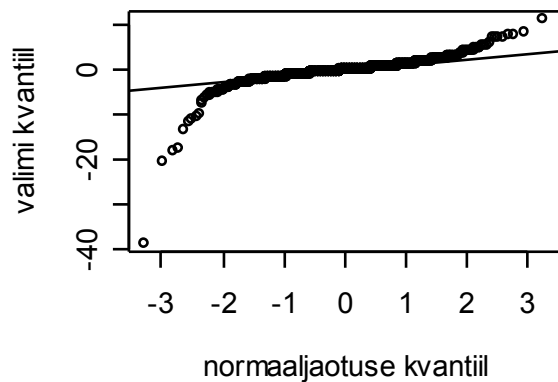
**pole normaaljaotus, n=100**



**normaaljaotus, n=1000**



**pole normaaljaotus, n=1000**



Kuidas joonistada tõenäosuspaberit tunnusele X?

Kasuta käske:

`qqnorm(X)` - joonistab punktid

`qqline(X)` - lisab pildile silma abistamiseks võrdlusjoone

Järgnevatele küsimustele vastates kontrolli enne, milliseid graafikuid võid saada, kui uuritav tunnus on tegelikult normaaljaotusega. Seda tee nii:

1. vaata, kui suur on valim (mitut vaatlust oled oma graafiku joonistamiseks kasutanud). Oletame järgmises näiteprogrammis, et vaatluseid oli 42.
2. tekita 42 vaatlusega valim garanteeritult normaaljaotusega uuritavast tunnusest. Normaaljaotusega populatsioonist võtab soovitud suurusega valimi R-i käsk `rnorm`. Joonista oma genereeritud andmete pealt graafik
3. korda sammu 2 mitu korda ja leia, kuivõrd kaugele „normist“ võib sellise suurusega valimi korral joonistav graafik jääda. Vaata, kas tegelike andmete pealt joonistatud graafik tundub veidram kui genereeritud andmete pealt joonistatud graafikud.

Seda tööd võiks teha järgmine programm. Programmi võiksid esmalt kirja panna tekstiredaktoris (notepad) ja sealt kogu programmi korruga kopeerida R-i.

Histogrammide joonistamine:

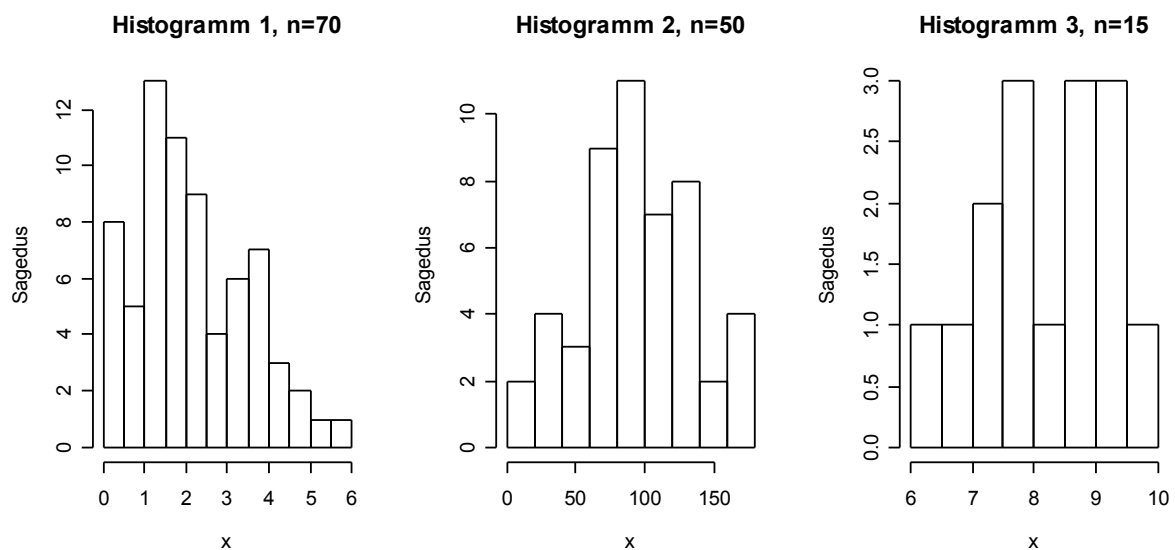
```
par(mfrow=c(2,2))
for (i in 1:4){
  x=rnorm(42)
  hist(x, main="normaaljaotus, n=42")
}
```

Tõenäosuspaberite joonistamine:

```
par(mfrow=c(2,2))
for (i in 1:4){
  x=rnorm(42)
  qqnorm(x, main="normaaljaotus, n=42")
  qqline(x)
}
```

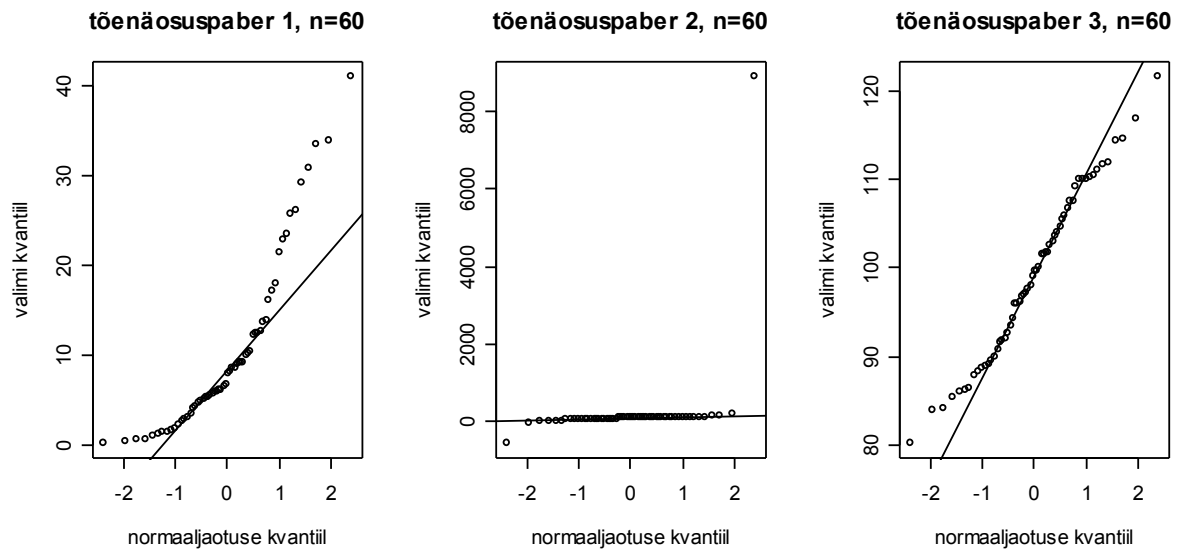
## Ülesanne 2

Otsusta, kas järgmiste graafikute puhul võiks olla tegemist normaaljaotusega tunnustega:



### Ülesanne 3

Vaata, millistel järgmistest graafikutest võiks olla uuritava tunnuse jaotuseks normaaljaotus?



### Ülesanne 4

Loe sisse eelmises praktikumis kasutatud kalade andmestik. Seda saab teha käsuga:

```
andmed=read.table("http://www.ms.ut.ee/mart/biomeetria2007/fishcatch.dat",  
header=TRUE)
```

Kontrolli, kas

- kalade kõrgus (*Height*) on normaaljaotusega;
- kas ahvenate (*Species=7*) kõrgus on normaaljaotusega;
- kas latikate (*Species=1*) kõrgus on normaaljaotusega;
- kas ahvenate kaal (*Weight*) on normaaljaotusega;