

Analyse af spruce datasættet

Benjamin Waziri

Februar 2024

Spruce datasættet

Datasættet er fra et eksperiment fra 90'erne, hvori en biolog plantede 72 træer under forskellige forhold (med/uden gødning, med/uden konkurrence), og målte træernes højde og diameter henover en periode på 5 år.

Data indlæses som følger:

```
library(resampledta3)

##
## Vedhæfter pakke: 'resampledta3'

## Det følgende objekt er maskeret fra 'package:datasets':
##
##      Titanic
```

a) Numeriske deskriptorer, højdevækst

Vi beregner, det opgaven beder os om:

```
summary(Spruce$Ht.change)
```

##	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
##	8.30	23.20	30.10	30.93	38.17	51.50

```
summary(Spruce$Di.change)
```

##	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
##	1.019	2.712	3.915	3.996	5.116	8.919

```
sd(Spruce$Ht.change)
```

```
## [1] 11.04943
```

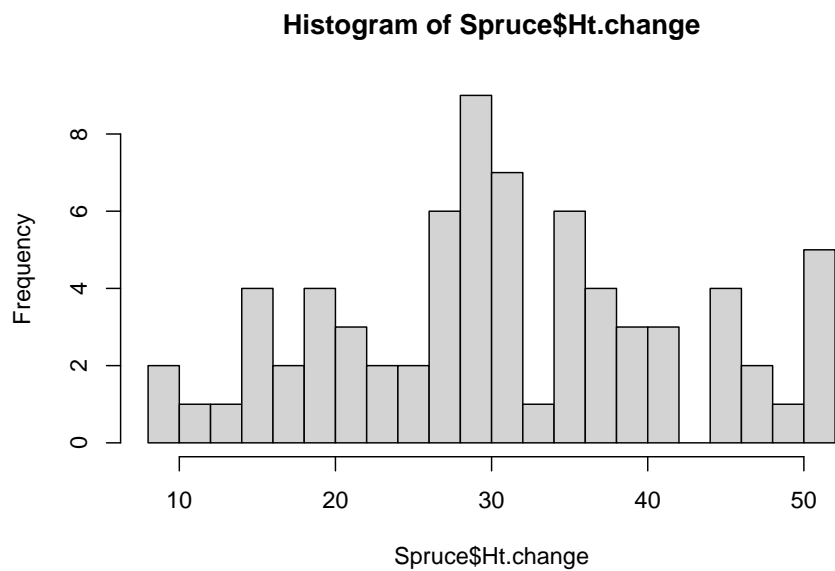
```
sd(Spruce$Di.change)
```

```
## [1] 1.787703
```

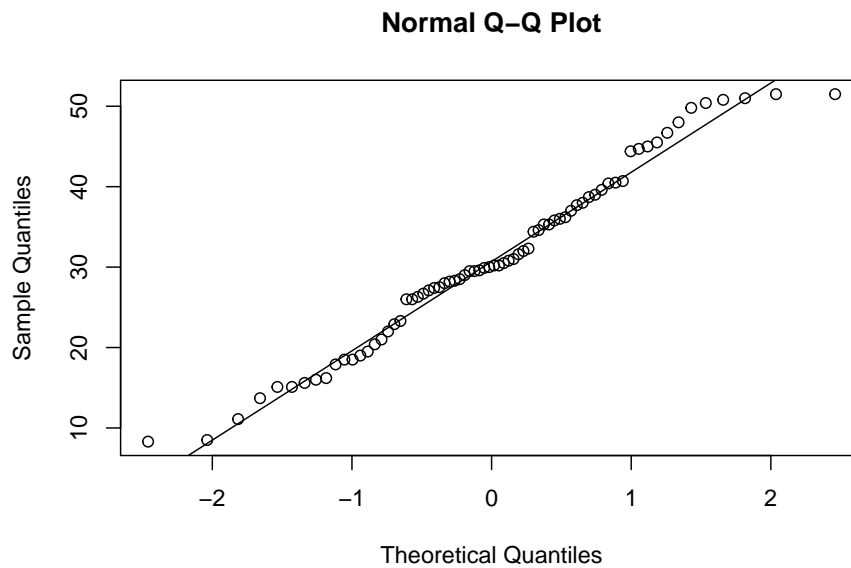
Fordelingen lader til at være relativt symmetrisk, da middelværdierne ligger tæt på medianerne.

b) Histogram og normalfraktilplot, højdevækst

```
hist(Spruce$Ht.change, breaks=30)
```



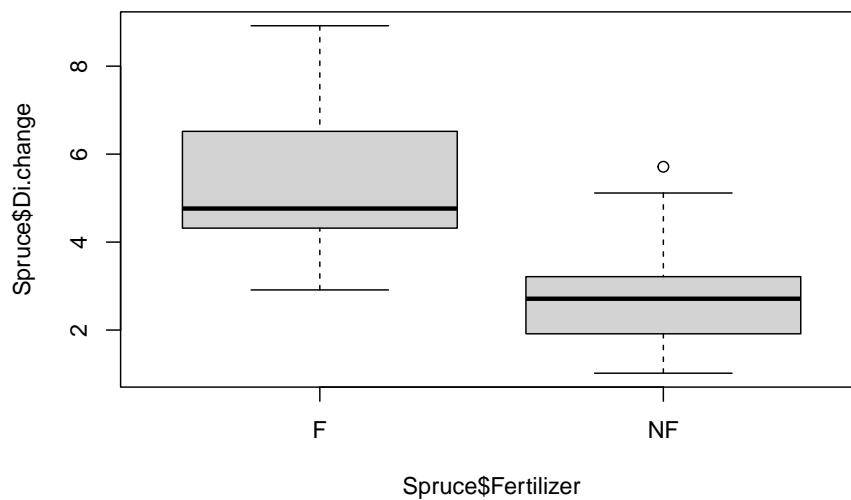
```
qqnorm(Spruce$Ht.change)  
qqline(Spruce$Ht.change)
```



Fordelingen er tilnærmelsesvis normal, da punkterne i vores QQ-plot tilnærmelsesvist ligger på en lige linje. Fordelingen afviger lidt fra en normalfordeling, dog er det grundet den lave sample size helt forventeligt.

c) Grafisk sammenligning af væksten i diameter, med og uden gødning

```
boxplot(Spruce$Di.change ~ Spruce$Fertilizer, data = Spruce$seedlings)
```



Her ses tydeligt, at gødning øger væksten. Der er en lille forskel i fordelingerne i form af, at medianen ligger meget tæt på Q1 i fordelingen for de træer, der modtog gødning. Det kunne tyde på en lidt skæv fordeling.

d) Numerisk sammenligning af væksten i diameter, med og uden gødning

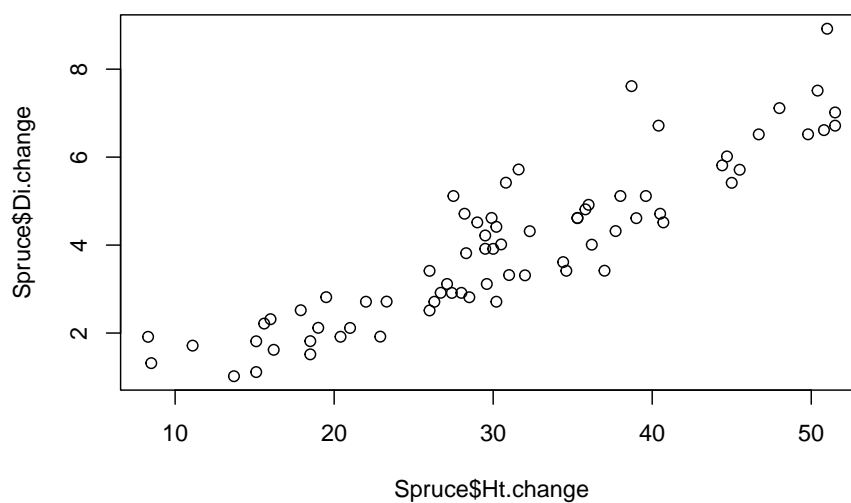
```
diameter_summary <- tapply(Spruce$Di.change, Spruce$Fertilizer, summary)
diameter_summary
```

```
## $F
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      2.913   4.318   4.763   5.274   6.518   8.919
##
## $NF
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      1.019   1.915   2.712   2.718   3.165   5.713
```

Her ses en langt større forskel på median og middelværdi hos de gødede træer end hos de ikke-gødede, hvilke understøtter, at de gødede træer har en lidt skæv fordeling.

e) Sammenhæng mellem væksten i højde og væksten i diameter

```
plot(Spruce$Ht.change, Spruce$Di.change)
```



Her ses en lineær positiv korrelation mellem forandringer i højde og forandringer i diameter.