

Figure 13-25. Left: QQ plot of height, which is close to normally distributed; right: QQ plot of age, which is not normally distributed

See Also

See `?qqplot` for information on comparing data to theoretical distributions other than the normal distribution.

`ggplot2` has a `stat_qq()` function, but it doesn't provide an easy way to draw the QQ line.

13.14. Creating a Graph of an Empirical Cumulative Distribution Function

Problem

You want to graph the empirical cumulative distribution function (ECDF) of a data set.

Solution

Use `stat_ecdf()` (Figure 13-26):

```
library(gcookbook) # For the data set
# ecdf of heightIn
```

```
ggplot(heightweight, aes(x=heightIn)) + stat_ecdf()

# ecdf of ageYear
ggplot(heightweight, aes(x=ageYear)) + stat_ecdf()
```

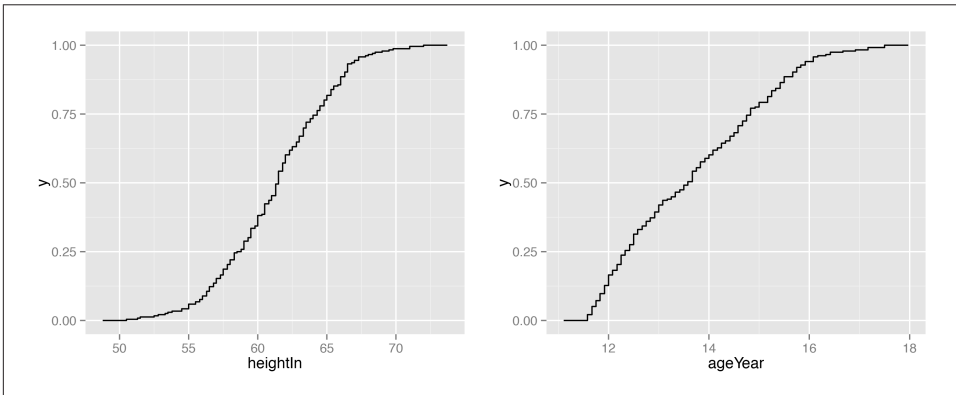


Figure 13-26. Left: ECDF of height; right: ECDF of age

Discussion

The ECDF shows what proportion of observations are at or below the given x value. Because it is *empirical*, the line takes a step up at each x value where there are one or more observations.

13.15. Creating a Mosaic Plot

Problem

You want to make a mosaic plot to visualize a contingency table.

Solution

Use the `mosaic()` function from the `vcd` package. For this example we'll use the `USBADmissions` data set, which is a contingency table with three dimensions. We'll first take a look at the data in a few different ways:

```
UCBADmissions
, , Dept = A

      Gender
Admit  Male Female
Admitted  512    89
Rejected  313    19
```