

# Statistical Inference: matched pairs and normal quantile plot

# Matched pairs

Some data:

| subject | druga | drugb |
|---------|-------|-------|
| 1       | 2.0   | 3.5   |
| 2       | 3.6   | 5.7   |
| 3       | 2.6   | 2.9   |
| 4       | 2.6   | 2.4   |
| 5       | 7.3   | 9.9   |
| 6       | 3.4   | 3.3   |
| 7       | 14.9  | 16.7  |
| 8       | 6.6   | 6.0   |
| 9       | 2.3   | 3.8   |
| 10      | 2.0   | 4.0   |
| 11      | 6.8   | 9.1   |
| 12      | 8.5   | 20.9  |

# Matched pairs data

- Data are comparison of 2 drugs for effectiveness at reducing pain.
  - 12 subjects (cases) were arthritis sufferers
  - Response is #hours of pain relief from each drug.
- In reading example, each child tried only one reading method.
- But here, each subject tried out both drugs, giving us two measurements.
  - Possible because, if you wait long enough, one drug has no influence over effect of other.
  - Advantage: focused comparison of drugs. Compare one drug with another on same person, removes a lot of variability due to differences between people.
  - Matched pairs, requires different analysis.
- Design: randomly choose 6 of 12 subjects to get drug A first, other 6 get drug B first.

## Paired t test: reading the data

Values aligned in columns:

```
my_url <- "http://www.utsc.utoronto.ca/~butler/c32/analgesic.t  
pain <- read_table(my_url)
```

```
## Parsed with column specification:  
## cols(  
##   subject = col_double(),  
##   druga = col_double(),  
##   drugb = col_double()  
## )
```

# The data

pain

| subject | druga | drugb |
|---------|-------|-------|
| 1       | 2.0   | 3.5   |
| 2       | 3.6   | 5.7   |
| 3       | 2.6   | 2.9   |
| 4       | 2.6   | 2.4   |
| 5       | 7.3   | 9.9   |
| 6       | 3.4   | 3.3   |
| 7       | 14.9  | 16.7  |
| 8       | 6.6   | 6.0   |
| 9       | 2.3   | 3.8   |
| 10      | 2.0   | 4.0   |
| 11      | 6.8   | 9.1   |
| 12      | 8.5   | 20.9  |

# Paired $t$ -test

```
with(pain, t.test(druga, drugb, paired = T))
```

```
##  
## Paired t-test  
##  
## data:  druga and drugb  
## t = -2.1677, df = 11, p-value = 0.05299  
## alternative hypothesis: true difference in means is not equal to  
## 95 percent confidence interval:  
## -4.29941513  0.03274847  
## sample estimates:  
## mean of the differences  
## -2.133333
```

- P-value is 0.053.
- Not quite evidence of difference between drugs.

## t-testing the differences

- Likewise, you can calculate the differences yourself and do a 1-sample t-test on them.
- First calculate a column of differences:

```
(pain %>% mutate(diff=druga-drugb) -> pain)
```

| subject | druga | drugb | diff  |
|---------|-------|-------|-------|
| 1       | 2.0   | 3.5   | -1.5  |
| 2       | 3.6   | 5.7   | -2.1  |
| 3       | 2.6   | 2.9   | -0.3  |
| 4       | 2.6   | 2.4   | 0.2   |
| 5       | 7.3   | 9.9   | -2.6  |
| 6       | 3.4   | 3.3   | 0.1   |
| 7       | 14.9  | 16.7  | -1.8  |
| 8       | 6.6   | 6.0   | 0.6   |
| 9       | 2.3   | 3.8   | -1.5  |
| 10      | 2.0   | 4.0   | -2.0  |
| 11      | 6.8   | 9.1   | -2.3  |
| 12      | 8.5   | 20.9  | -12.4 |

## t-test on the differences

- then throw them into `t.test`, testing that the mean is zero, with same result as before:

```
with(pain,t.test(diff,mu=0))
```

```
##  
## One Sample t-test  
##  
## data: diff  
## t = -2.1677, df = 11, p-value = 0.05299  
## alternative hypothesis: true mean is not equal to 0  
## 95 percent confidence interval:  
## -4.29941513 0.03274847  
## sample estimates:  
## mean of x  
## -2.133333
```



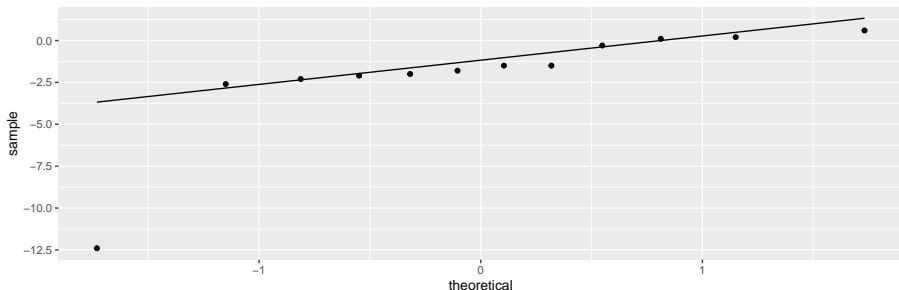
# Assessing normality

- 1-sample and 2-sample t-tests assume (each) group normally distributed.
- Matched pairs analyses assume (theoretically) that differences normally distributed.
- Though we know that t-tests generally behave well even without normality.
- How to assess normality? A normal quantile plot.
  - Idea: scatter of points should follow the straight line, without curving.
  - Outliers show up at bottom left or top right of plot as points off the line.

# The normal quantile plot

- of differences from matched pairs data

```
ggplot(pain, aes(sample=diff)) + stat_qq() + stat_qq_line()
```



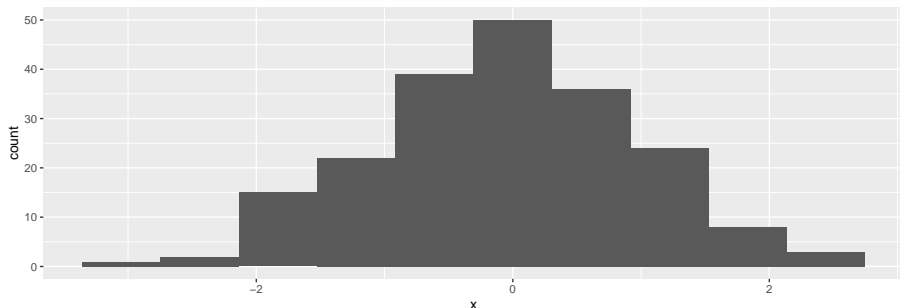
- Points should follow the straight line. Bottom left one way off, so normality questionable here: outlier.

# More normal quantile plots

- How straight does a normal quantile plot have to be?
- There is randomness in real data, so even a normal quantile plot from normal data won't look perfectly straight.
- With a small sample, can look not very straight even from normal data.
- Looking for systematic departure from a straight line; random wiggles ought not to concern us.
- Look at some examples where we know the answer, so that we can see what to expect.

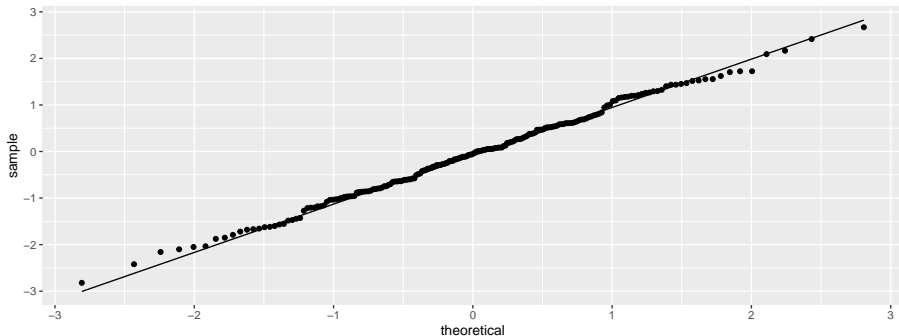
# Normal data, large sample

```
d=tibble(x=rnorm(200))  
ggplot(d,aes(x=x))+geom_histogram(bins=10)
```



# The normal quantile plot

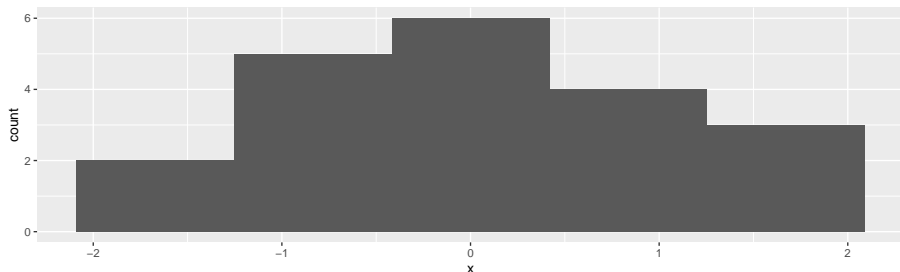
```
ggplot(d,aes(sample=x))+stat_qq()+stat_qq_line()
```



# Normal data, small sample

- Not so convincingly normal, but not obviously skewed:

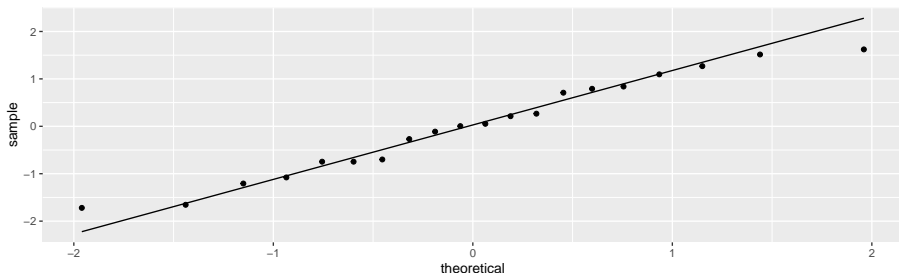
```
d=tibble(x=rnorm(20))  
ggplot(d,aes(x=x))+geom_histogram(bins=5)
```



# The normal quantile plot

Good, apart from the highest and lowest points being slightly off. I'd call this good:

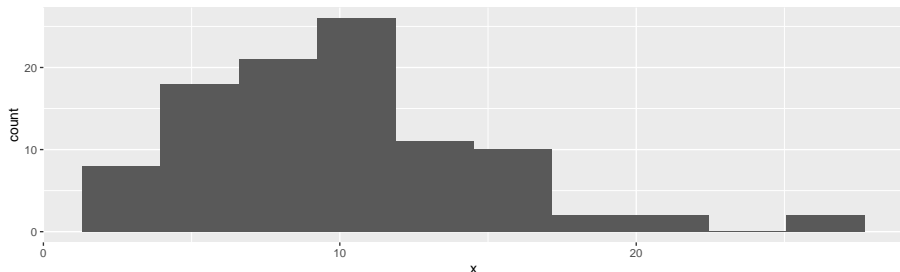
```
ggplot(d, aes(sample=x)) + stat_qq() + stat_qq_line()
```



## Chi-squared data, $df = 10$

Somewhat skewed to right:

```
d=tibble(x=rchisq(100,10))  
ggplot(d,aes(x=x))+geom_histogram(bins=10)
```

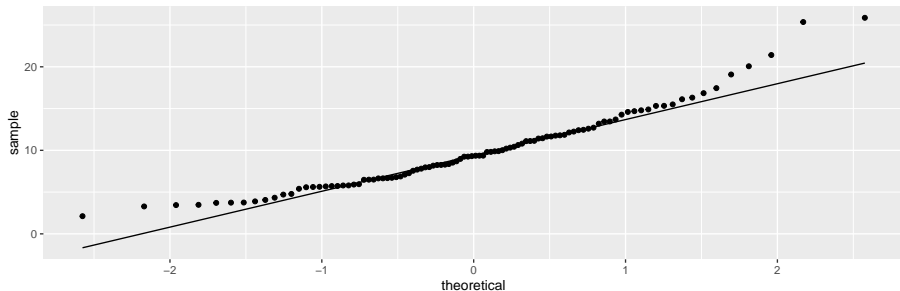




# The normal quantile plot

Somewhat opening-up curve:

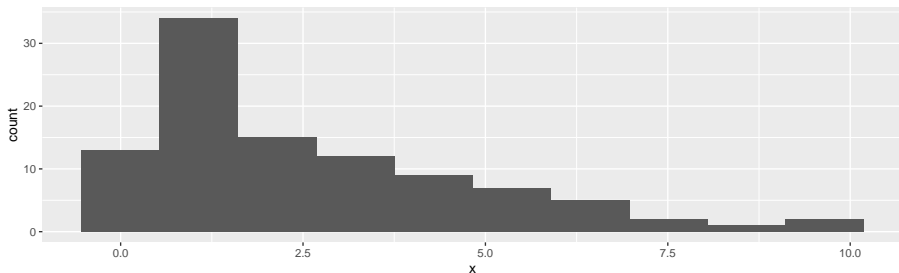
```
ggplot(d,aes(sample=x))+stat_qq()+stat_qq_line()
```



## Chi-squared data, $df = 3$

Definitely skewed to right:

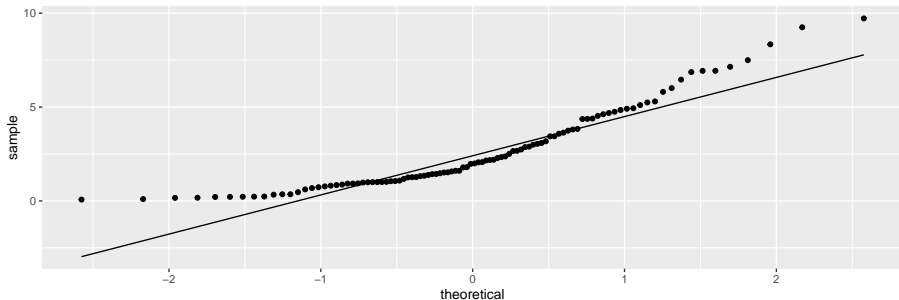
```
d=tibble(x=rchisq(100,3))  
ggplot(d,aes(x=x))+geom_histogram(bins=10)
```



# The normal quantile plot

Clear upward-opening curve:

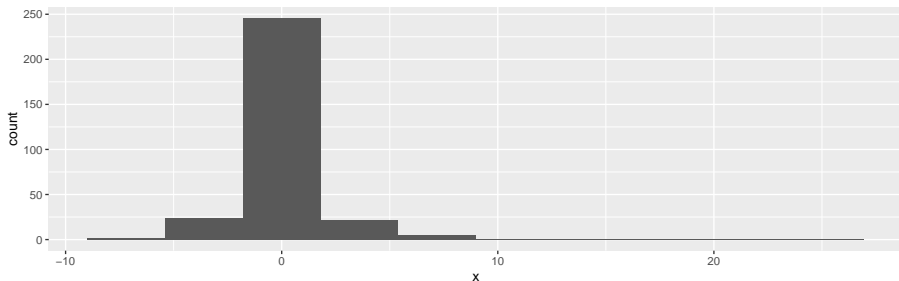
```
ggplot(d,aes(sample=x))+stat_qq()+stat_qq_line()
```



## t-distributed data, $df = 3$

Long tails (or a very sharp peak):

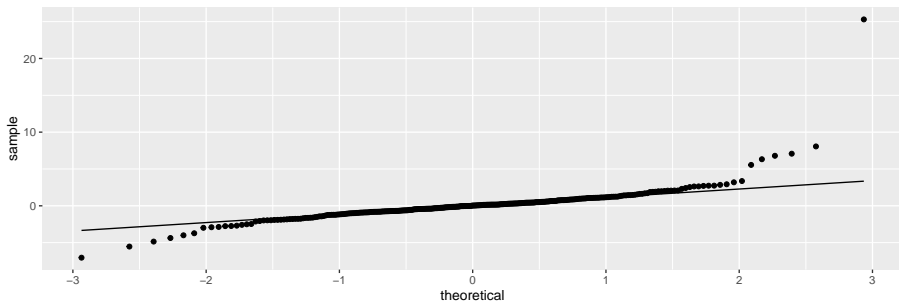
```
d=tibble(x=rt(300,3))  
ggplot(d,aes(x=x))+geom_histogram(bins=10)
```



# The normal quantile plot

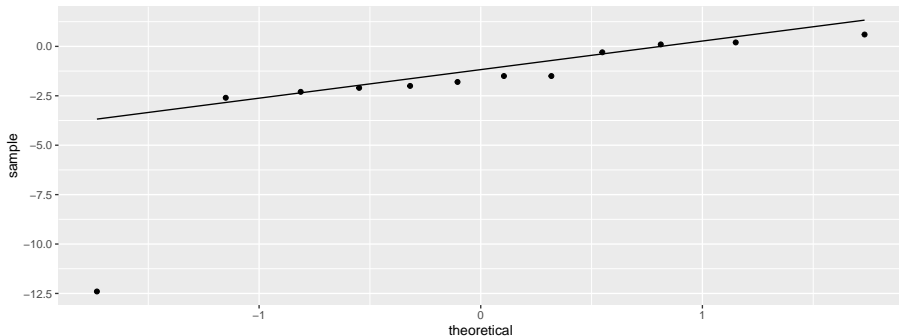
Low values too low and high values too high for normal.

```
ggplot(d,aes(sample=x))+stat_qq()+stat_qq_line()
```



# Our pain-relief data

```
ggplot(pain, aes(sample=diff)) + stat_qq() + stat_qq_line()
```



# Comments

- Definitely not normal. What to do?
- Sign test on differences, null median 0.

# Sign test

- Most easily: calculate differences in data frame, then use `smmr`.
- Null median difference is 0:

```
pain %>% mutate(mydiff=druga-drugb) %>%  
sign_test(mydiff,0)
```

```
## $above_below  
## below above  
##      9      3  
##  
## $p_values  
##   alternative    p_value  
## 1         lower 0.07299805  
## 2          upper 0.98071289  
## 3    two-sided 0.14599609
```



# Comments

- P-value 0.1460. No evidence that the drugs are different.
- Since we are working in a pipeline, input data frame to `sign_test` is “whatever came out of previous step”.

## (Some of) the kids' reading data, again

```
## Parsed with column specification:  
## cols(  
##   group = col_character(),  
##   score = col_double()  
## )
```

```
kids %>% slice_sample(n=12)
```

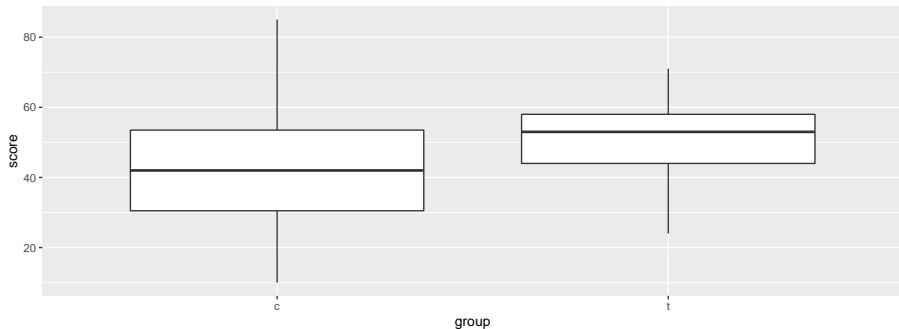
| group | score |
|-------|-------|
| t     | 49    |
| c     | 48    |
| c     | 46    |
| c     | 17    |
| t     | 43    |
| t     | 44    |
| c     | 26    |
| .     | 50    |

# Where we were at

- 21 kids in “treatment”, new reading method; 23 in “control”, standard reading method.
- Assessing assumptions:
  - We did two-sample t-test (Satterthwaite-Welch) before.
  - Assumes approx. normal data within each group.
  - Does not assume equal spread.
  - (Pooled t-test *does* assume equal spread).
  - Assess each group separately.

# Boxplots for reading data

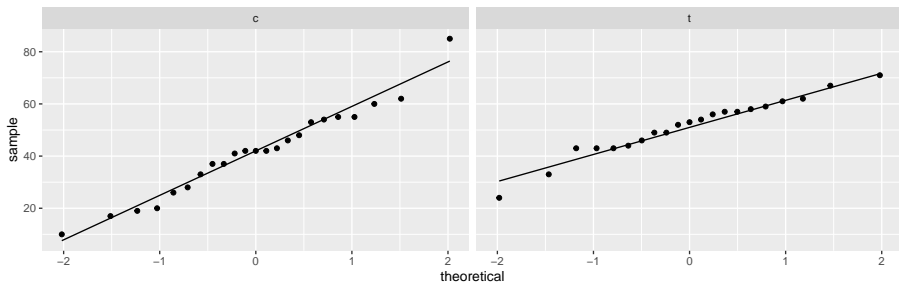
```
ggplot(kids, aes(x=group, y=score)) + geom_boxplot()
```



# Facetted normal quantile plots

Done this way:

```
ggplot(kids, aes(sample=score)) + stat_qq() + stat_qq_line() +  
facet_wrap(~group)
```



# Comments

- These plots show no problems with normality. Both groups are more or less symmetric/normal and there are no outliers.
- Equal spreads questionable, but we don't need that.
- Assess equal spreads by looking at *slopes* of normal quantile plots.
- We ought be happy with the (Welch) two-sample t-test (over)

# Welch two-sample test

```
t.test(score~group,data=kids,alternative="less")
```

```
##  
##  Welch Two Sample t-test  
##  
## data:  score by group  
## t = -2.3109, df = 37.855, p-value = 0.01319  
## alternative hypothesis: true difference in means is less than 0  
## 95 percent confidence interval:  
##      -Inf -2.691293  
## sample estimates:  
## mean in group c mean in group t  
##      41.52174      51.47619
```

from which we concluded that the new reading method really does help.