One-Sample t-Test

```
> t.test(rnorm(20,mean=2,sd=2))

One Sample t-test

data:  rnorm(20, mean = 2, sd = 2)
t = 2.384, df = 19, p-value = 0.02771
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 0.1558692 2.3985594
sample estimates:
mean of x
 1.277214

> t.test(rnorm(20,mean=0,sd=2))

One Sample t-test

data:  rnorm(20, mean = 0, sd = 2)
t = -0.9111, df = 19, p-value = 0.3737
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -1.5884003  0.6249483
sample estimates:
mean of x
-0.481726

>
```
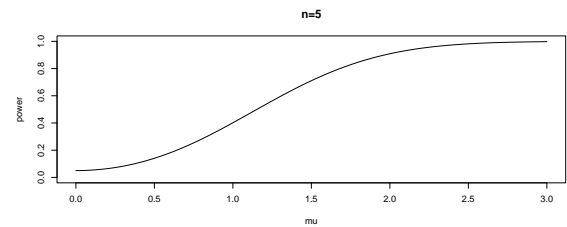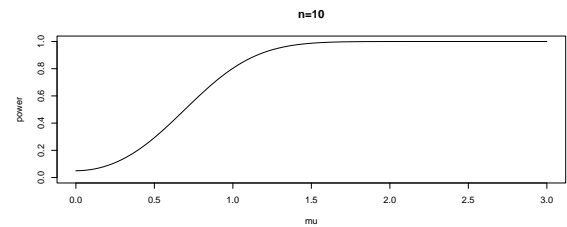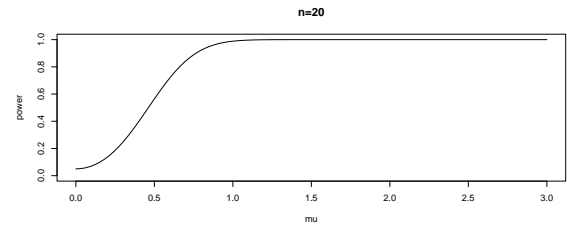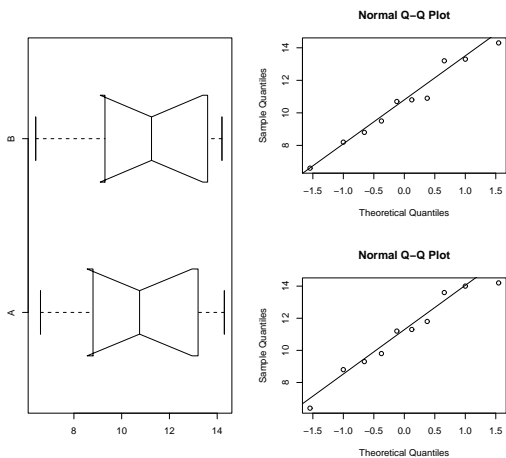
Power Curves for t-test

$\alpha = .05,\ X_1, \ldots, X_n \sim N(\mu, 1)$

Shoe Wear

```
> library(MASS)
> shoes
$A
 [1] 13.2  8.2 10.9 14.3 10.7  6.6  9.5 10.8  8.8 13.3
$B
 [1] 14.0  8.8 11.2 14.2 11.8  6.4  9.8 11.3  9.3 13.6
> layout(rbind(c(1,2),c(1,3)))
> boxplot(list(A=shoes$A, B=shoes$B,
          notch=T, horizontal=T, boxwex=.5)
> qqnorm(shoes$A); qqline(shoes$A)
> qqnorm(shoes$B); qqline(shoes$B)
```

Paired t-Test

```
> t.test(shoes$A, shoes$B, paired=T)

Paired t-test

data:  shoes$A and shoes$B
t = -3.3489, df = 9, p-value = 0.008539
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.6869539 -0.1330461
sample estimates:
mean of the differences
               -0.41

> t.test(shoes$A-shoes$B)

One Sample t-test

data:  shoes$A - shoes$B
t = -3.3489, df = 9, p-value = 0.008539
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.6869539 -0.1330461
sample estimates:
mean of x
    -0.41

>
```

Paired vs. Unpaired

```
> t.test(shoes$A, shoes$B, paired=T)

Paired t-test

data:  shoes$A and shoes$B
t = -3.3489, df = 9, p-value = 0.008539
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.6869539 -0.1330461
sample estimates:
mean of the differences
                  -0.41

> t.test(shoes$A, shoes$B)  # bad idea: using paired=F

Welch Two Sample t-test

data:  shoes$A and shoes$B
t = -0.3689, df = 17.987, p-value = 0.7165
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -2.745046  1.925046
sample estimates:
mean of x mean of y
    10.63     11.04

>
```
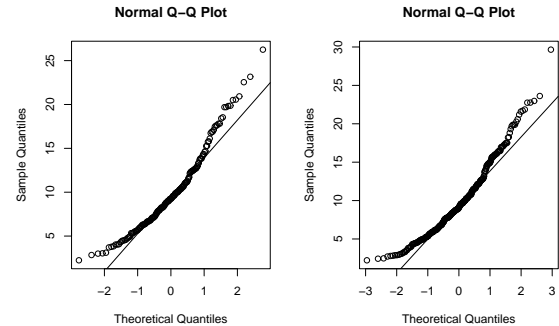
Non-normality of Traffic Data

```
> library(MASS)
> attach(Traffic)
> qqnorm(y[limit=="yes"]); qqline(y[limit=="yes"])
> qqnorm(y[limit=="no"]); qqline(y[limit=="no"])
> shapiro.test(y[limit=="yes"])
Shapiro-Wilk normality test
data:  y[limit == "yes"]
W = 0.9213, p-value = 0.0003330
> shapiro.test(y[limit=="no"])
Shapiro-Wilk normality test
data:  y[limit == "no"]
W = 0.9516, p-value = 0.0003928
>
```


Normal Q–Q Plot   **Normal Q–Q Plot**

Normality (?) of Log Traffic Data

```
> qqnorm(log(y[limit=="yes"]))
> qqnorm(log(y[limit=="no"]))
> shapiro.test(log(y[limit=="yes"]))
Shapiro-Wilk normality test
data:  log(y[limit == "yes"])
W = 0.9832, p-value = 0.4814
> shapiro.test(log(y[limit=="no"]))
Shapiro-Wilk normality test
data:  log(y[limit == "no"])
W = 0.9868, p-value = 0.3235
>
```


Normal Q–Q Plot   **Normal Q–Q Plot**

Two-Sample t-Test

```
> ly.yes <- log(y[limit=="yes"])
> ly.no <- log(y[limit=="no"])
> t.test(ly.yes, ly.no)
Welch Two Sample t-test
data:  ly.yes and ly.no
t = -3.2954, df = 147.673, p-value = 0.001231
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.31472006 -0.07876154
sample estimates:
mean of x mean of y
 2.866770  3.063511
> var.test(ly.yes, ly.no)
F test to compare two variances
data:  ly.yes and ly.no
F = 0.9262, num df = 68, denom df = 114, p-value = 0.7389
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.6109935 1.4372433
sample estimates:
ratio of variances
         0.9262063
> t.test(ly.yes, ly.no, var.equal=T)
Two Sample t-test
data:  ly.yes and ly.no
t = -3.2638, df = 182, p-value = 0.001313
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.31567636 -0.07780524
sample estimates:
mean of x mean of y
 2.866770  3.063511
>
```
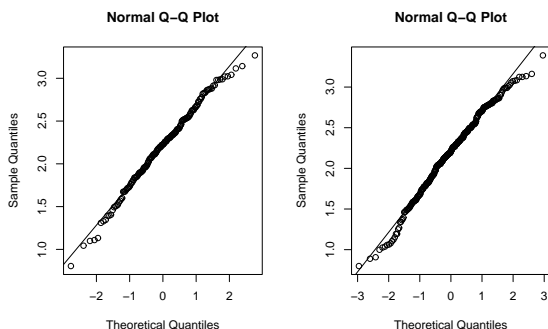
Approximate Z-Tests in R

Instead of log-transforming traffic data, let's do an approximate Z-test on the untransformed data:

- Same statistic, so use same test function:
  ```
  > t.test(y[limit=="yes"],y[limit=="no"])

   Welch Two Sample t-test

  data:  y[limit == "yes"] and y[limit == "no"]
  t = -3.3995, df = 165.545, p-value = 0.000846
  alternative hypothesis: true difference
  in means is not equal to 0
  95 percent confidence interval:
   -6.666816 -1.767967
  sample estimates:
  mean of x mean of y
   18.91304   23.13043
  ```

- If you want to be picky, use p-value based on Z instead of $t(165.545)$:
  ```
  > 2*pt(-3.3995,df=165.545,lower.tail=T)
  [1] 0.0008459695
  > 2*pnorm(-3.3995,lower.tail=T)
  [1] 0.0006750918
  >
  ```

  but this is all so approximate anyway, why not take the more conservative t-based p-value?

Recall that "exact" p-value for log-transformed data assuming normality and equal variances was $p = 0.001313$.

Wilcoxon Signed-Rank Test

- For one sample:
  ```
  > x <- c(8.5,8.6,6.4,12.1,8.2,7.4,7.8,8.3,10.3,8.4)
  > wilcox.test(x, mu=10, conf.int=T)
  Wilcoxon signed rank test
  data:  x
  V = 8, p-value = 0.04883
  alternative hypothesis: true mu is not equal to 10
  95 percent confidence interval:
   7.50 9.95
  sample estimates:
  (pseudo)median
          8.35
  ```

- For paired data:
  ```
  > wilcox.test(shoes$A, shoes$B, paired=T)
  Wilcoxon signed rank test with continuity correction
  data:  shoes$A and shoes$B
  V = 3, p-value = 0.01431
  alternative hypothesis: true mu is not equal to 0
  Warning message: Cannot compute exact p-value with ties in:
  wilcox.test.default(shoes$A, shoes$B, paired = T)
  > wilcox.test(shoes$A-shoes$B)
  Wilcoxon signed rank test with continuity correction
  data:  shoes$A - shoes$B
  V = 3, p-value = 0.01431
  alternative hypothesis: true mu is not equal to 0
  Warning message: Cannot compute exact p-value with ties in:
  wilcox.test.default(shoes$A - shoes$B)
  >
  ```
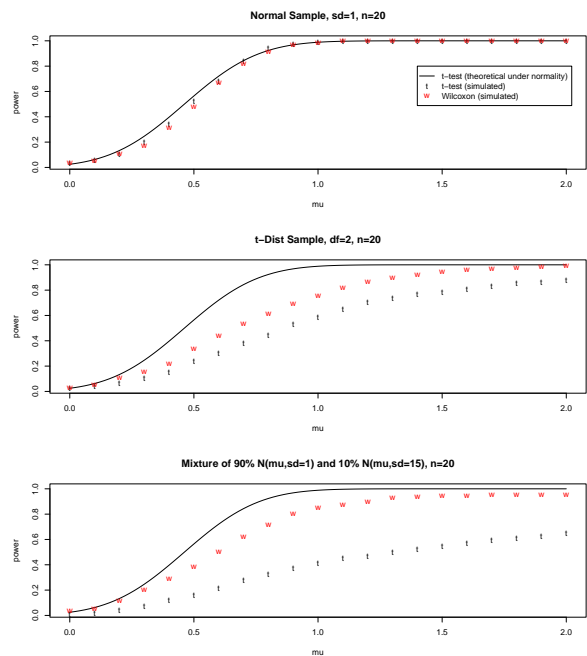
Why I Love the Wilcoxon Test
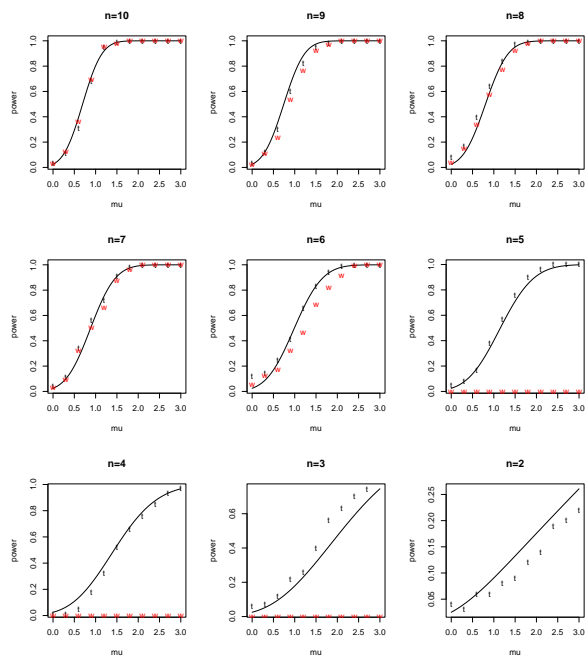
Why I Love It Even More

Two-Sample Wilcoxon Test

```
> wilcox.test(y.yes,y.no)

Wilcoxon rank sum test with continuity correction

data:  y.yes and y.no
W = 2878, p-value = 0.001830
alternative hypothesis: true mu is not equal to 0

> wilcox.test(log(y.yes),log(y.no))

Wilcoxon rank sum test with continuity correction

data:  log(y.yes) and log(y.no)
W = 2878, p-value = 0.001830
alternative hypothesis: true mu is not equal to 0

>
```
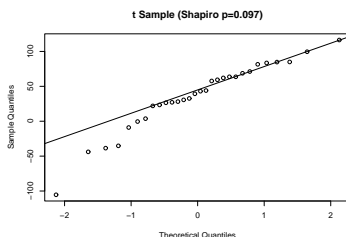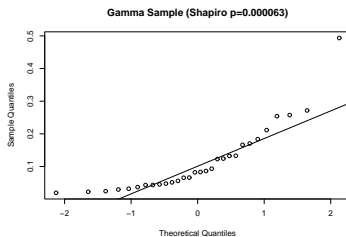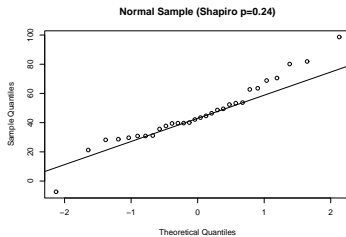
Power of Shapiro Test

```
> rejects <- 0
> for (i in 1:10000) {
+   x <- rnorm(30, mean=45, sd=20)
+   if (shapiro.test(x)$p.value < 0.05) rejects <- rejects + 1
+ }
> rejects
[1] 530
> rejects <- 0
> for (i in 1:10000) {
+   x <- rgamma(30, shape=2, rate=20)
+   if (shapiro.test(x)$p.value < 0.05) rejects <- rejects + 1
+ }
> rejects
[1] 7513
> rejects <- 0
> for (i in 1:10000) {
+   x <- 30+45*rt(30, df=4)
+   if (shapiro.test(x)$p.value < 0.05) rejects <- rejects + 1
+ }
> rejects
[1] 3249
>
```

Power of Shapiro Test

Normal Sample (Shapiro p=0.24)

Gamma Sample (Shapiro p=0.000063)

t Sample (Shapiro p=0.097)

Multigroup Location Tests

```
> names(PlantGrowth)
[1] "weight" "group"
> levels(PlantGrowth$group)
[1] "ctrl" "trt1" "trt2"
> oneway.test(weight ~ group, data=PlantGrowth)

One-way analysis of means (not assuming equal variances)

data:  weight and group
F = 5.181, num df = 2.000, denom df = 17.128, p-value = 0.01739

> oneway.test(weight ~ group, data=PlantGrowth, var.equal=T)

One-way analysis of means

data:  weight and group
F = 4.8461, num df = 2, denom df = 27, p-value = 0.01591

> summary(aov(weight ~ group, data=PlantGrowth))
            Df  Sum Sq Mean Sq F value  Pr(>F)
group        2  3.7663  1.8832  4.8461 0.01591 *
Residuals   27 10.4921  0.3886
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> kruskal.test(weight ~ group, data=PlantGrowth)

Kruskal-Wallis rank sum test

data:  weight by group
Kruskal-Wallis chi-squared = 7.9882, df = 2, p-value = 0.01842

>
```

Testing Two-Way Contingencies

```
> table(r,stat404)
   stat404
r  - n  y
 0 2 4  0
 1 3 4  1
 2 0 7 14
 3 1 1 10
> chisq.test(table(r,stat404))

Pearson's Chi-squared test

data:  table(r, stat404)
X-squared = 21.9431, df = 6, p-value = 0.00124

Warning message:
Chi-squared approximation may be incorrect
in: chisq.test(table(r, stat404))
> fisher.test(table(r,stat404))

Fisher's Exact Test for Count Data

data:  table(r, stat404)
p-value = 0.0001139
alternative hypothesis: two.sided
>
```

Testing Two-Way Contingencies

```
> table(matlab,stat404)
      stat404
matlab - n  y
     0 5 9 17
     1 0 6  5
     2 1 1  1
     3 0 0  2
> chisq.test(table(matlab,stat404))

Pearson's Chi-squared test

data:  table(matlab, stat404)
X-squared = 6.3824, df = 6, p-value = 0.3817

Warning message:
Chi-squared approximation may be incorrect
in: chisq.test(table(matlab, stat404))
> fisher.test(table(matlab,stat404))

Fisher's Exact Test for Count Data

data:  table(matlab, stat404)
p-value = 0.3786
alternative hypothesis: two.sided

>
```

Bigger Example

```
> library(MASS)
> names(Melanoma)
[1] "time"      "status"    "sex"       "age"       "year"
[6] "thickness" "ulcer"
> attach(Melanoma)
> table(sex,ulcer)
   ulcer
sex 0  1
  0 79 47
  1 36 43
> chisq.test(table(sex,ulcer))

Pearson's Chi-squared test with Yates' continuity correction

data:  table(sex, ulcer)
X-squared = 5.1099, df = 1, p-value = 0.02379

> fisher.test(table(sex,ulcer))

Fisher's Exact Test for Count Data

data:  table(sex, ulcer)
p-value = 0.02061
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
 1.089871 3.700065
sample estimates:
odds ratio
  2.000701


>
```
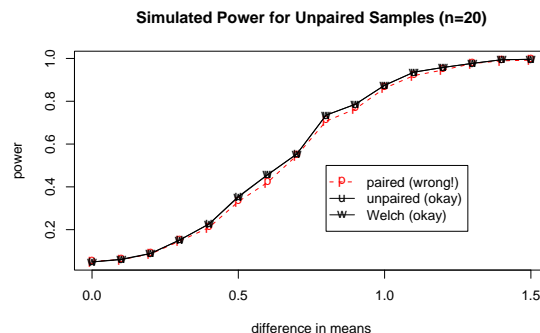
For Unpaired (Independent) Samples

Independent, iid samples:

$$X_1, \ldots, X_{20} \sim N(\mu_X, 1)$$
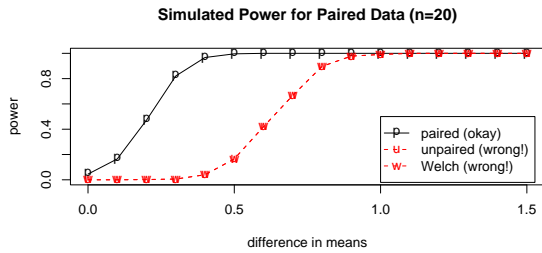$$Y_1, \ldots, Y_{20} \sim N(\mu_Y, 1)$$

At $\alpha = 0.05$,

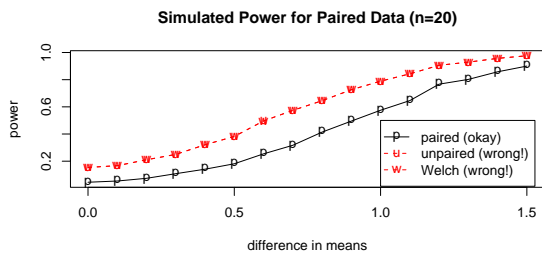**Simulated Power for Unpaired Samples (n=20)**

For Paired Data

$(X_1, Y_1), \ldots, (X_{20}, Y_{20})$ iid bivariate normal.

At $\alpha = 0.05$,

- Positive correlation (usual) $\rho = +0.9$

Simulated Power for Paired Data (n=20)

- Negative correlation (unusual) $\rho = -0.9$

Simulated Power for Paired Data (n=20)

Summary of Tests

Location tests:

- `t.test`: t-tests for one sample, paired, two-sample (equal variance and Welch's approximation). Can also be used for approximate Z-tests.
- `wilcox.test`: Wilcoxon signed-rank test for one sample or paired data, Wilcoxon/Mann-Whitney rank-sum test for two samples.
- `oneway.test`: F-test for difference in means of two or more samples (equal variance and Welch's approximation).
- `kruskal.test`: Kruskal-Wallis rank-sum test for difference in means of two or more samples.

Categorical data tests:

- `chisq.test`: χ^2-test for one-way and interaction in two-way contingency tables.
- `fisher.test`: Fisher's exact test for interaction in two-way contingency tables.

Other tests:

- `shapiro.test`: Shapiro-Wilk test for H_0 : sample is normal.
- `var.test`: F-test for H_0 : samples have equal variances.