Nonparametric tests

- “Distribution free” methods require fewer assumptions than parametric methods
- Focus on testing rather than estimation
- Not sensitive to outlying observations
- Especially useful for cruder data (like ranks)
- “Throws away” some of the information in the data
- May be less powerful than parametric counterparts, when the parametric assumptions are true
- For large samples, are equally efficient to parametric counterparts

Fish	SR	P	Diff	Sgn rank	Fish	SR	P	Diff	Sng rank
1	.32	.39	.07	+15.5	13	.20	.22	.02	+6.5
2	.40	.47	.07	+15.5	14	.31	.30	-.01	−2.5
3	.11	.11	.00		15	.62	.60	-.02	−6.5
4	.47	.43	-.04	−11.0	16	.52	.53	.01	+2.5
5	.32	.42	.10	+20.0	17	.77	.85	.08	+17.5
6	.35	.30	-.05	−13.5	18	.23	.21	-.02	−6.5
7	.32	.43	.11	+20.0	19	.30	.33	.03	+9.0
8	.63	.98	.35	+23.0	20	.70	.57	-.13	−21.0
9	.50	.86	.36	+24.0	21	.41	.43	.02	+6.5
10	.60	.79	.19	+22.0	22	.53	.49	-.04	−11.0
11	.38	.33	-.05	−13.5	23	.19	.20	.01	+2.5
12	.46	.45	-.01	−2.5	24	.31	.35	.04	+11.0
					25	.48	.40	-.08	−17.5

Measurements are mecury levels in fish (ppm)

Alternatives to the paired t-test

- Let $D_i =$ difference (`P - SR`)
- Let θ be the population median of the D_i
- $H_0 : \theta = 0$ versus $H_a : \theta \neq 0$ (or $>$ or $<$)
- Notice that $\theta = 0$ iff $p = P(D > 0) = .5$
- Let X be the number of times $D > 0$

  X is then binomial(n, p)
- The sign test tests wether $H_0 : p = .5$ using X

Example

$\theta =$ median difference `p - sr`

$H_0 : \theta = 0$ versus $H_a : \theta \neq 0$

Number of instances where the difference is bigger than 0 is 15 out of 25 trials

```
binom.test(15, 25)
```

```
p-value = 0.4244
```

Or we could have used large sample tests for a binomial proportion `prop.test(15, 25, p = .5)`

```
X-squared = 0.64, df = 1, p-value = 0.4237
```

Discussion

- Magnitude of the differences is discarded

  Perhaps too much information lost

- Could easily have tested $H_0 : \theta = \theta_0$ by calculating the number of times $D > \theta_0$ and performing a binomial test

  We can invert these tests to get a distribution free confidence interval for the median

Signed rank test

- Wilcoxon's statistic uses the information in the **signed ranks** of the differences
- Saves some of the information regarding the magnitude of the differences
- Still tests $H_0 : \theta = 0$ versus the three alternatives
- Appropriately normalized, the test statistic follows a normal distribution
- Also the exact small sample distribution of the signed rank statistic is known (if there are no ties)

Signed rank procedure

1. Take the paired differences
2. Take the absolute values of the differences
3. Rank these absolute values, throwing out the 0s
4. Multiply the ranks by the sign of the difference (+1 for a positive difference and -1 for a negative difference)
5. Cacluate the rank sum W_+ of the positive ranks

Signed rank procedure

- If $\theta > 0$ then W_+ should be large
- If $\theta < 0$ then W_+ should be small
- Properly normalized, W_+ follows a large sample normal distribution
- For small sample sizes, W_+ has an exact distribution under the null hypothesis
  * Can get critical values from tables in the textbook
  * How could you use Monte Carlo to calculate an exact P-value (if there are no ties)?

Large sample distribution of W_+

- Under H_0 and if there are no ties

$$E(W_+) = n(n+1)/4$$

$$Var(W_+) = n(n+1)(2n+1)/24$$

$$TS = \{W_+ - E(W_+)\}/Sd(W_+) \to \mathrm{Normal}(0,1)$$

- There is a correction term necessary for ties
- Without ties, it's possible to do an exact (small sample) test

Example

```
diff <- c(.07, .07, .00, -.04,  ...)
wilcox.text(diff, exact = FALSE)
```

- H_0 : Med diff $= 0$ vesus H_a : Med diff $\neq 0$
- $W_+ = 194.5$
- $E(W_+) = 24 \times 25/4 = 150$
- $\text{Var}(W_+) = 24 \times 25 \times 49/24 = 1,225$
- $TS = (194.5 - 150)/\sqrt{1,224} = 1.27$
- P-value $= .20$
- R's P-value (uses correction for ties) $= 0.21$

Methods for unpaired samples

Comparing two measuring techniques A and B

Units are in deg C per gram

Method A	Method B
79.98 80.05	80.02
80.04 80.03	79.94
80.02 80.02	79.98
80.04 80.00	79.97
80.03 80.02	79.97
80.03	80.03
80.04	79.95
79.97	79.97

The Mann-Whitney test

- Tests whether or not the two treatments have the same location
- Assumes independent identically distributed errors, not necessarily normal
- Null hypothesis can also be written more generally as a stochastic shift for two arbitrary distributions
- Test uses the sum of the ranks obtained by discarding the treatment lables
- Also called the Wilcoxon rank sum test

The Mann-Whitney test

- Procedure

1. Discard the treatment labels
2. Rank the observations
3. Calculate the sum of the ranks in the first treatment
4. Either
   * calculate the asymptotic normal distrubtion of this statistic
   * compare with the exact distribution under the null hypothesis

Method A		Method B
7.5	21.0	11.5
19.0	15.5	1.0
11.5	11.5	7.5
19.0	9.0	4.5
15.5	11.5	4.5
15.5		15.5
19.0		2.0
4.5		4.5
180		51

Sum has to add up to $21 \times 22/2 = 231$

Aside

Gauss supposedly came up with this in grade school

```
x =   1   +    2   +    3   +    4   +   ...   +    n
x =   n   +  n-1   +  n-2   +  n-3   +   ...   +    1
```

Therefore

```
2x = n+1  +  n+1  +  n+1  +  n+1  +  ...  +  n+1
```

So 2x = n (n + 1) / 2

So x = n (n + 1) / 2

Results

Let W be the sum of the ranks for the first treatment (A)
Let n_A and n_B be the sample sizes
Then

- $E(W) = n_A(n_a + n_b + 1)/2$
- $\mathrm{Var}(W) = n_A n_B(n_A + n_B + 1)/12$
- $TS = \{W - E(W)\}/Sd(W) \to N(0,1)$

Also the exact distribution of W can be calculated

Example

- $W = 51$
- $E(W) = 8(8 + 13 + 1)/2 = 88$
- $Sd(W) = \sqrt{8 \times 13(8 + 13 + 1)/12} = 13.8$
- $TS = (51 - 88)/13.8 = -2.68$
- Two-sided P-value$= .007$
- R function `wilcox.test` will perform the test

Notes about nonpar tests

- Tend to be more robust to outliers than parametric counterparts
- Do not require normality assumptions
- Usually have exact small-sample versions
- Are often based on ranks rather than the raw data
- Loss in power over parametric counterparts is often not bad
- Nonpar tests are not assumption free