STAT 500

Blocking and Matched Pairs

Key to Statistical Significance

- Differences on response variable between groups are larger than the differences within groups.
  - t-test statistic

$$t = \frac{(\bar{Y}_i - \bar{Y}_j) - 0}{S_p\sqrt{\frac{1}{n_i} + \frac{1}{n_j}}}$$

  - F-test statistic

$$F = \frac{MS_{\text{model}}}{MS_{\text{error}}}$$

Variation within Groups

- When σ^2 is large compared to differences between means
  - Fail to reject H_0 of equal means even when differences between means exist.
- Why would σ^2 be large?
  - Response variable has large amount of variation.
  - Experimental units are not homogeneous with respect to response variable.

Variation within Groups

- Choose more homogeneous experimental units.
  - Reduces variation in response variable - more likely to produce significant result.
  - Reduces generalizability of experimental results.
- Use more heterogeneous experimental units.
  - Increases variation in response variable - less likely to produce significant result.
  - Increases generalizability of experimental results.

Reduce Variation through Blocking

- Block: a group of experimental units that, prior to treatment, are expected to be more like one another (with respect to response variables) than experimental units in general.
  In simple words, blocks are groups of similar experimental units.

- Group experimental units into blocks before assigning treatments

- Randomly assign treatments to experimental units within each block separately.

Blocks through Sorting

You are interested in the effect of two different instructional methods on achievement in mathematics of 8th graders. Sort students by their Iowa Test math scores from 7th grade. Students within each block will have similar Iowa Test math scores.

Blocks through Subdividing

You are interested in the yield of three varieties of soy beans. You have 12 fields across Iowa that you can use. Divide each field into 3 sections and plant one variety on each section.

Blocks through Reusing

You are interested in the determining which of two brands of golf balls travels the furthest when hit with a five iron. Have each person hit both types of golf ball (reuse each person).

Blocks through Matching

You are interested in the determining which of two brands of golf balls travels the furthest when hit with a five iron. Pair two golfers with same skill level. Have one person hit one brand of golf ball and other person hit the other brand of golf ball.

Effect of Blocking

- Reduce $\boldsymbol{\sigma}^{\mathbf{2}} =$ variation within experimental units
- Variation due to block variable is removed from $\boldsymbol{\sigma}^{\mathbf{2}}$ (and its estimate)
- Easier to detect difference between treatment groups

Matched Pairs

- Experiments with two treatments
- Blocks have one or two experimental units
  - One unit (reuse)
    - Receives both treatments
    - Order of treatments is random
  - Two units (match)
    - Two treatments randomly assigned to pair
    - One unit receives one treatment - other unit receives other treatment

Matched Pairs Data

- Y_{i1} = response to treatment 1 in the ith block
- Y_{i2} = response to treatment 2 in the ith block
- n blocks, assumed independent

	Treatment 1	Treatment 2
Block 1	Y_{11}	Y_{12}
Block 2	Y_{21}	Y_{22}
$\vdots$	$\vdots$	$\vdots$
$\vdots$	$\vdots$	$\vdots$
$\vdots$	$\vdots$	$\vdots$
Block n	Y_{n1}	Y_{n2}

Matched Pairs Data

- Responses within a block are not independent; usually $\boldsymbol{Y_{i1}}$ has a positive correlation with $\boldsymbol{Y_{i2}}$

- $\boldsymbol{\mu_1}$ = mean response for treatment 1

- $\boldsymbol{\mu_2}$ = mean response for treatment 2

- $\boldsymbol{\mu_d = \mu_1 - \mu_2}$ = mean difference in responses between treatments

Analysis of Matched Pairs Data

- $D_i = Y_{i1} - Y_{i2}$ = difference in response between treatments for each of the n blocks

- Estimate μ_d with

$$\bar{D} = \frac{1}{n}\sum_{i=1}^{n} D_i = \frac{1}{n}\sum_{i=1}^{n}(Y_{i1} - Y_{i2}) = \bar{Y}_1 - \bar{Y}_2$$

Analysis of Matched Pairs Data

$$
\begin{aligned}
Var(\bar{D}) &= Var(\frac{1}{n}\sum_{i=1}^{n} D_i) \\
&= \frac{1}{n^2} Var(\sum_{i=1}^{n} D_i) \\
&= \frac{1}{n^2}\sum_{i=1}^{n} Var(D_i) \quad \text{(assumes independent blocks)} \\
&= \frac{1}{n^2}\sum_{i=1}^{n} Var(Y_{i1} - Y_{i2}) \\
&= \frac{1}{n^2}\sum_{i=1}^{n} [Var(Y_{i1}) + Var(Y_{i2}) - 2Cov(Y_{i1}, Y_{i2})]
\end{aligned}
$$

Analysis of Matched Pairs Data

$$
\begin{aligned}
Var(\bar{D}) &= \frac{1}{n^2}\sum_{i=1}^{n}\left[Var(Y_{i1}) + Var(Y_{i2}) - 2Cov(Y_{i1}, Y_{i2})\right] \\
&= \frac{1}{n^2}\sum_{i=1}^{n}\left[\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2\right] \\
&= \frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{n} - 2\frac{\rho\sigma_1\sigma_2}{n} \\
&= Var(\bar{Y}_1) + Var(\bar{Y}_2) - 2Cov(\bar{Y}_1, \bar{Y}_2)
\end{aligned}
$$

- ρ is *within* block correlation between Y_{i1} and Y_{i2}
- ρ is positive, making overall $Var(\bar{D})$ smaller

Analysis of Matched Pairs Data

Estimate $Var(\bar{D})$ using observed differences D_i

- Unbiased estimate of $Var(D_i)$

$$s_d^2 = \frac{1}{n-1} \sum_{i=1}^{n} (D_i - \bar{D})^2$$

- Standard Error of $\bar{D}$

$$\text{SE}(\bar{D}) = \frac{s_d}{\sqrt{n}}$$

Analysis of Matched Pairs Data

- Hypothesis Test
  - $H_0 : \mu_d = 0$ vs. $H_a : \mu_d \neq 0$
  - Test Statistic:
    $$t = \frac{\bar{D}}{(s_d/\sqrt{n})}$$
  - P-value:
    $$2 \times P(t_{n-1} > |t|)$$

Analysis of Matched Pairs Data

- $100(1-\alpha)\%$ Confidence Interval for μ_d

$$\bar{D} \pm t_{n-1,1-\alpha/2}\frac{s_d}{\sqrt{n}}$$

Monkey Nerve Cells Study

Example: Creatine phosphate (CP) concentrations in rhesus monkey nerve cells.

- n=8 monkeys
- Nerves extending from one side of the spinal cord were severed
- CP concentrations (mg per 100g of tissue) were measured during the regeneration process
- Nerves extending from the other side of the spinal cord were kept intact (control).
- CP concentrations were also measured on the control side.

Monkey Nerve Cells Study

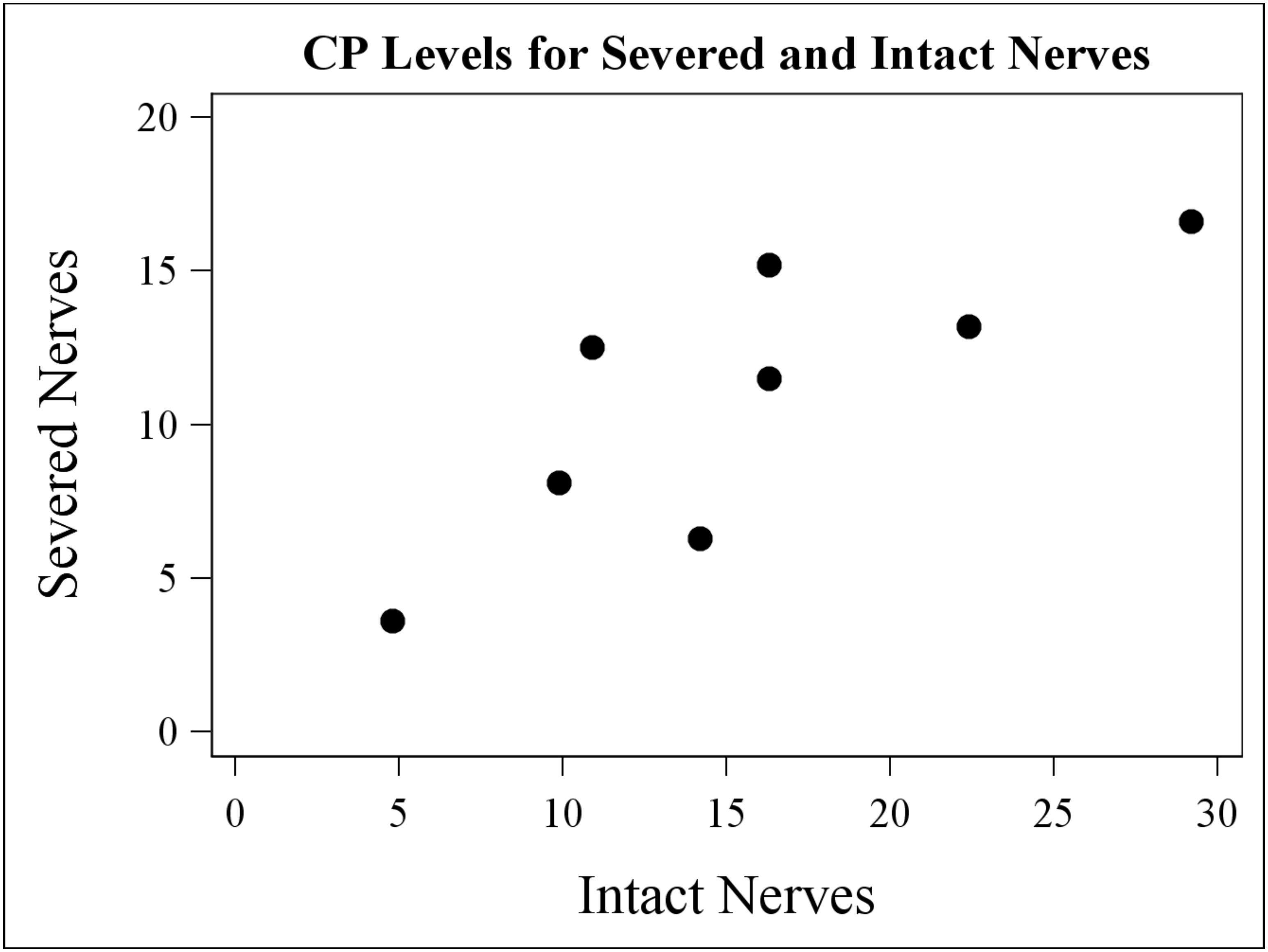
Example: Creatine phosphate (CP) concentrations in rhesus monkey nerve cells.

- Randomization
  - Randomly select four monkeys to have nerves severed on the right side
  - The other four monkeys will have nerves severed on the left side

Monkey Nerve Cells Study

Example: Creatine phosphate (CP) concentrations in rhesus monkey nerve cells.

Monkey	Severed nerves	Intact nerves
1	11.5	16.3
2	3.6	4.8
3	12.5	10.9
4	6.3	14.2
5	15.2	16.3
6	8.1	9.9
7	16.6	29.2
8	13.1	22.4


CP Levels for Severed and Intact Nerves
Severed Nerves
Intact Nerves
0
5
10
15
20
0
5
10
15
20
25
30

Monkey Nerve Cells Study

- Experimental units: monkeys
- Treatments: severed nerve, control
- Paired data (repeated measurements)
  - One side (severed nerves)
  - Other side (control)
- Used a paired t-test
  - Test $H_o: \mu_{severed} = \mu_{control}$
  - against $H_a: \mu_{severed} \neq \mu_{control}$

Monkey Nerve Cell Study

Pair	Difference (severed-control)
1	$d_1 = Y_{11} - Y_{12} = -4.8$
2	$d_2 = Y_{21} - Y_{22} = -1.0$
3	$d_3 = Y_{31} - Y_{32} = 1.6$
4	$d_4 = Y_{41} - Y_{42} = -7.9$
5	$d_5 = Y_{51} - Y_{52} = -1.1$
6	$d_6 = Y_{61} - Y_{62} = -1.8$
7	$d_7 = Y_{71} - Y_{72} = -12.6$
8	$d_8 = Y_{81} - Y_{82} = -9.3$

- Sample mean: $\bar{d} = -4.64$
- Sample variance: $S_d^2 = 23.87$
- Standard error for mean difference: $S_{\bar{d}} = \sqrt{\frac{S_d^2}{n}} = 1.73$

Monkey Nerve Cell Study

- t-statistic:

$$t = \frac{\bar{d} - 0}{S_{\bar{d}}} = -2.685 \text{ on } n - 1 = 7 \text{ df}$$

- Two-sided p-value=.0313

- Using a type I error level of $\alpha = 0.05$ we can reject the null hypothesis and conclude that CP concentration is different during nerve cell regeneration from normal tissue.

- A 95% confidence interval for the difference in mean CP concentrations for severed and intake nerve cells

$$\bar{d} \pm t_{(n-1),1-\alpha/2}\sqrt{\frac{S_d^2}{n}}$$

Alternative Computation of the Sample Variance of the Differences

$$Var(\bar{D}) = \frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{n} - 2\frac{\rho\sigma_1\sigma_2}{n}$$

is also estimated as

$$\begin{aligned} S_{\bar{D}}^2 &= \frac{S_1^2}{n} + \frac{S_2^2}{n} - 2\frac{\hat{\rho}S_1\ S_2}{n} \\ &= \frac{20.21}{8} + \frac{57.90}{8} - 2\frac{(.79469)(4.4956)(7.609)}{8} \\ &= 2.97 \end{aligned}$$

$$S_{\bar{d}} = \sqrt{2.97} = 1.72$$

Consequence of Doing the Wrong Test

If the pairing (correlation) is incorrectly ignored we would compute

$$S_{\bar{Y}_1-\bar{Y}_2} = \sqrt{\frac{S_1^2}{n} + \frac{S_2^2}{n}}$$

$$= \sqrt{\frac{20.2107}{8} + \frac{57.8971}{8}} = 3.06$$

and

$$t = \frac{\bar{Y}_1 - \bar{Y}_2}{S_{\bar{Y}_1-\bar{Y}_2}} = -1.48 \text{ on } 11.4 \text{ df}$$

with two-sided p-value=.1660

Paired Analysis - Diagnostics

- Model assumptions are
  - Blocks are independent $\rightarrow$ differences are independent
  - D_i are i.i.d. $N(\mu_d = \mu_1 - \mu_2, \sigma_d^2)$

    where $\sigma_d^2 = \sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2$

Paired Analysis - Diagnostics

- Independence of differences
  - Examine study to determine if responses from one block could affect responses from any other block.
  - Critical problem if this fails
  - Observations from same block will usually be positively correlated

Paired Analysis - Diagnostics

- Normal distribution for differences
  - Normal probability plot for differences
  - Effects of non-normality
    - t-test sensitive to outliers
    - t-test sensitive to skewness of the distribution of possible differences
    - If sample size is large, t-test is fairly robust to these problems

Nonparametric Tests

- If smaller sample sizes with non-normal differences
  - Wilcoxon signed rank test
  - Sign test

Nonparametric Tests

- Wilcoxon signed rank test
  - Order the absolute values of the differences
  - Assign ranks to the differences
  - Test statistic: T = sum of ranks corresponding to positive differences
  - T is approximately normal (for large n) with mean $n(n+1)/4$ and variance $n(n+1)(2n+1)/24$
  - Corrections for tied values (see any text on nonparametric statistics).
  - Not quite as good as a t-test for normal data, but much better than a t-test for skewed distributions

Monkey Nerve Cell Study

Pair	$\lvert Difference \rvert$	rank
1	$\lvert d_1 \rvert = \lvert -4.8 \rvert = 4.8$	5
2	$\lvert d_2 \rvert = \lvert -1.0 \rvert = 1.0$	1
3	$\lvert d_3 \rvert = \lvert 1.6 \rvert = 1.6$	3
4	$\lvert d_4 \rvert = \lvert -7.9 \rvert = 7.9$	6
5	$\lvert d_5 \rvert = \lvert -1.1 \rvert = 1.1$	2
6	$\lvert d_6 \rvert = \lvert -1.8 \rvert = 1.8$	4
7	$\lvert d_7 \rvert = \lvert -12.6 \rvert = 12.6$	8
8	$\lvert d_8 \rvert = \lvert -9.3 \rvert = 9.3$	7

- T = 3 and n=8
- $|Z| = \frac{|T - n(n+1)/4| - 0.5}{\sqrt{n(n+1)(2n+1)/24}} = \mathbf{2.11}$ with p-value=$\mathbf{0.035}$
- Exact p-value=0.039

Nonparametric Tests

- Sign test
  - Based on number of positive and number of negative differences.
  - Ties (zero differences) are ignored
  - Ignores the sizes of the observed differences
  - Not as powerful as Wilcoxon signed rank
  - Can be useful for censored data
  - p-value obtained from the binomial distribution (use normal approximation to binomial for large samples)

Monkey Nerve Cell Study

Pair	Difference	sign
1	$\lvert d_1 \rvert = -4.8$	-
2	$\lvert d_2 \rvert = -1.0$	-
3	$\lvert d_3 \rvert = 1.6$	$+$
4	$\lvert d_4 \rvert = -7.9$	-
5	$\lvert d_5 \rvert = -1.1$	-
6	$\lvert d_6 \rvert = -1.8$	-
7	$\lvert d_7 \rvert = -12.6$	-
8	$\lvert d_8 \rvert = -9.3$	-

- S = 1 positive difference out of n=8 pairs
- $H_o : Pr(+) = Pr(-) = 0.5$ vs $H_A : Pr(+) \neq 0.5$
- p-value $= 2\left[\binom{8}{0}(0.5)^8 + \binom{8}{1}(0.5)^8\right] = 0.070$