STUDY DESIGNS

IN BIOMEDICAL RESEARCH

CROSS-OVER DESIGNS

A confounder is a factor (a subject's characteristic) which may be related to the treatment or the outcome even the factor itself may not be under investigation. A study may involve one or several confounders.

In theory, values of confounders may have been balanced out between study groups because patients were randomized. But it is not guaranteed; especially if the sample size is not very large.

If confounder or confounders are known, heterogeneous experimental units are divided into homogeneous "block"; and randomizations of treatments are carried out within each block (stratified randomization). The result would be a "randomized complete block design".

Sometimes we are not aware of all confounders; and some confounders are hard to quantify. And post hoc solution (i.e. regression) requires assumptions which may not hold. Therefore, it is more desirable to handle confounders by a proper experiment design.

THE ROLE OF STUDY DESIGN

In a "standard" experimental design, a linear model for a continuous response/outcome is:

$$Y = \left[\begin{array}{c}\text{Overall}\\ \text{Constant}\end{array}\right] + \left[\begin{array}{c}\text{Treatment}\\ \text{Effect}\end{array}\right] + \left[\begin{array}{c}\text{Experimental}\\ \text{Error}\end{array}\right]$$

The last component, "experimental error", includes not only error specific to the experimental process but also includes "subject effect" (age, gender, etc…). Sometimes these subject effects are large making it difficult to assess "treatment effect".

Blocking (to turn a completely randomized design into a randomized complete block design) would help. But it would only help to “reduce” subject effects, not to “eliminate” them: subjects in the same block are only similar, not identical – unless we have “blocks of size one”. And that the basic idea of “Cross-over Design”, a very popular form in biomedical research.

Cross-over is a very special design where we have “bloc” of size one; each subject serves as his/her own control receiving both treatment. Randomization decides the order. The outcome could be binary or on the continuous scale.

Let start with the case of a continuous outcome and, for illustration, consider a project we just completed here: a clinical trial to prevent lung cancers.

We designed and conducted a placebo-controlled cross-over clinical trial to assess the effect of a PEITC supplement on changes of NNK metabolism in smokers.

THE DESIGN

In the following "PEITC trial", measurements (urinary total NNAL) will be taken from each subject in the two supplementation sequences as seen in the following diagram:

Group 1: Period #1 (PEITC; A1) – washout – Period #2 (Placebo; B2)
Group 2: Period #1 (Placebo; B1) – washout – Period #2 (PEITC; A2)

The letter is used to denote supplementation or treatment (A for PEITC and B for Placebo) and the number, 1 or 2, denotes the period; e.g. "A1" for PEITC taken in period #1.

The “washout periods” are inserted in order to eliminate possible “carry-over effects” (The half-life of dietary PEITC in vivo is between 2-3 hours, with complete excretion within 1-2 days following ingestion).

REGRESSION MODELS

Group 1: Period #1 (PEITC; A1) – washout – Period #2 (Placebo; B2)
Group 2: Period #1 (Placebo; B1) – washout – Period #2 (PEITC; A2)

The mean of A1, A2, B1, and B2 can be modeled as follows:

$$\text{Mean} = (\alpha)(\text{Treatment}) + (\beta)(\text{Order}) + \text{Others}$$

Treatment is coded as (0 = Placebo, 1 =PEITC)
Order is coded as (0= 2nd Period, 1= 1st Period)

Others include all subjects' characteristics

OUTCOME VARIABLES

From the design:

Group 1: Period #1 (PEITC; A1) – washout – Period #2 (Placebo; B2)

Group 2: Period #1 (Placebo; B1) – washout – Period #2 (PEITC; A2)

Our data analysis could be based on the following "outcome variables" (Treatment - Placebo):

X1 = A1 - B2; and X2 = A2 - B1

This subtraction will cancel the "within-sequence" effects of all subject-specific factors. This process will result in two independent samples (often with the same or similar sample size if there are no or minimal dropouts or missing data).

Recall the general model:

$$Y = \left[\begin{array}{c}\text{Overall}\\ \text{Constant}\end{array}\right] + \left[\begin{array}{c}\text{Treatment}\\ \text{Effect}\end{array}\right] + \left[\begin{array}{c}\text{Experimental}\\ \text{Error}\end{array}\right]$$

The subtractions

X1 = A1 - B2; and X2 = A2 - B1

will cancel not only effects of all subject-specific factors; they cancel the "overall constant" as well, leaving only two parameters in the means of X1 and X2:

Mean of X1 = [(α)(1)+(β)(1)] – [(α)(0)+(β)(0)] = α + β

Mean of X2 = [(α)(1)+(β)(0)] – [(α)(0)+(β)(1)] = α - β

RESULTING LINEAR MODELS

Design:

Group 1: Period #1 (PEITC; A1) – washout – Period #2 (Placebo; B2)

Group 2: Period #1 (Placebo; B1) – washout – Period #2 (PEITC; A2)

Outcome variables:

X1 = A1 - B2; and X2 = A2 - B1

Resulting Linear regression models:

X1 is normally distributed as $N(\alpha+\beta,\sigma^2)$

X2 is normally distributed as $N(\alpha-\beta,\sigma^2)$

In this models, α represents the PEITC supplementation effect (α>0 if and only if PEITC increases the total NNAL) and β represents the period effect (β>0 if and only if measurement from period 1 is larger than from period 2).

(1) The X's do not really need to have normal distributions; the robustness comes from the fact that our analysis will be based on the normal distribution of the sample mean – not of the data, and the sample mean would be almost normally distribution for moderate to large sample sizes (Central Limit Theorem).

(2) Among the three parameters, α represents the PEITC effect and is the primary target, β could be of some interest; we have no interest in σ^2 (we have to handle it properly to make inferences on α (and β) valid and efficient.

DATA ANALYSIS

Design:

Group 1: Period #1 (PEITC; A1) – washout – Period #2 (Placebo; B2)

Group 2: Period #1 (Placebo; B1) – washout – Period #2 (PEITC; A2)

Outcome variables:

X1 = A1 - B2; and X2 = A2 - B1

From the model:

X1 is normally distributed as $N(\alpha+\beta,\sigma^2)$

X2 is normally distributed as $N(\alpha-\beta,\sigma^2)$

Let the sample means and sample variances be defined as usual and n the group size (total sample size is 2n); then, we can easily prove the followings:

TREATMENT EFFECTS

X1 is normally distributed as $N(a+\beta,\sigma^2)$
X2 is normally distributed as $N(a-\beta,\sigma^2)$

$$\bar{x}_1 \text{ is } distributed \text{ as normal } N(\alpha+\beta, \frac{\sigma^2}{n}),$$

$$\bar{x}_2 \text{ is } distributed \text{ as normal } N(\alpha-\beta, \frac{\sigma^2}{n})$$

$$\mathbf{a} = \frac{\bar{\mathbf{x}}_1 + \bar{\mathbf{x}}_2}{\mathbf{2}}$$

$$\mathbf{a \text{ is } distributed \text{ as normal } N(\alpha, \frac{\sigma^2}{2n})}$$

ESTIMATION OF PARAMETERS

(1) Estimation of Variance: We can pool data from the two sequences to estimate the common variance σ^2 by s_p^2 – the same pooled estimate used in two-sample t-test.

(2) Estimation of Treatment Effect: Parameter α representing the PEITC effect, the difference between PEITC and the placebo, is estimated by a – the average of the two sample means. Its 95 percent confidence interval is given by:

$$a \pm t_{.975} \frac{s_p}{\sqrt{N}}$$

The t-coefficient goes with (N-2) degrees of freedom; without missing data, N = 2n – total number of subjects.

TESTING FOR TREATMENT EFFECT

Testing for PEITC Treatment Effect: Null hypothesis of "no treatment effects" H0: $\alpha = 0$ is tested using the "t test", with (N-2) degrees of freedom:

$$t = \frac{a}{s_p / \sqrt{N}}$$

It's kind of "one-sample t-test" but we use the degree of freedom associated with s_p. Alternatively, one can frame it as a two-sample t-test comparing the mean of X1 versus the mean of (-X2) as seen from

X1 is normally distributed as $N(\alpha+\beta,\sigma^2)$

X2 is normally distributed as $N(\alpha-\beta,\sigma^2)$

ORDER EFFECTS

X1 is normally distributed as $N(\alpha+\beta,\sigma^2)$
X2 is normally distributed as $N(\alpha-\beta,\sigma^2)$

$$\bar{x}_1 \text{ is } distributed \text{ as normal } N(\alpha+\beta, \frac{\sigma^2}{n}),$$

$$\bar{x}_2 \text{ is } distributed \text{ as normal } N(\alpha-\beta, \frac{\sigma^2}{n})$$

$$\mathbf{b} = \frac{\bar{x}_1 - \bar{x}_2}{2}$$

$$\mathbf{b \text{ is } distributed \text{ as normal } N(\beta, \frac{\sigma^2}{2n})}$$

ESTIMATION OF PARAMETERS

(1) Estimation of Variance: We can pool data from the two sequences to estimate the common variance σ^2 by s_p^2 – the same pooled estimate used in two-sample t-test.

(2) Estimation of Order Effect: Parameter β representing the order effect, the difference between Period 1 and Period 2, is estimated by b – half the difference of the two sample means. Its 95 percent confidence interval is given by:

$$b \pm t_{.975} \frac{s_p}{\sqrt{N}}$$

The t-coefficient goes with (N-2) degrees of freedom; without missing data, N = 2n – total number of subjects.

TESTING FOR ORDER EFFECT

Testing for Order Effect: Null hypothesis of "no order effects"

H0: $\beta = 0$ is tested using the "t test", with (N-2) degrees of freedom:

$$t = \frac{b}{s_p / \sqrt{N}}$$

It's kind of "one-sample t-test" but we use the degree of freedom associated with s_p. Alternatively, one can frame it as a two-sample t-test comparing the mean of X1 versus the mean of X2 as seen from

X1 is normally distributed as $N(\alpha+\beta,\sigma^2)$

X2 is normally distributed as $N(\alpha-\beta,\sigma^2)$

In crossover designs, one could make one single observation per period, or one could make multiple observations over different time points in the same period. If there are multiple observations per period, we could average these multiple observations, and us the simple method as just presented.

For crossover studies, the treatment effects from the earlier period might be carried over to the later period, which is called "carry-over effect". Due to this possible effect, the response might be changed over different period, so it could be referred to as an "interaction". We could not investigate it by the simple regression model.

To deal with these problems, we have a more powerful tool, the "linear mixed-effects" (LME) model. In this use, we model separately the treatment effects, the period effects, and the carryover effects or interaction

LME MODEL FOR CROSS-OVER DESIGN

MODEL:

Define two indicator variables representing the Treatment (TRT) and Period (PRD) effects:

TRT = 1 for Treatment A and TRT = -1 for Treatment B

PRD = 1 for Period 1 and PRD = -1 for Period 2

And let X_{ijk} denote the response/observation for ith sequence, jth subject in kth period.

We consider this **LME Model for Crossover Design**:

$$X_{ijk} = \beta_0 + \beta_1 TRT + \beta_2 PRD + \beta_{12}(TRT)(PRD) + s_{ij} + \varepsilon_{ijk}$$

$$X_{ijk} = \beta_0 + \beta_1 TRT + \beta_2 PRD + \beta_{12}(TRT)(PRD) + s_{ij} + \varepsilon_{ijk}$$

The β‘s are fixed effects whereas s_{ij} is a term representing the random subject effects (a random intercept); s_{ij} is assumed to be independently and identically distributed as “normal” with mean zero, $N(0,\sigma_s^2)$. The last term, ε_{ijk}, denotes the error which is assumed to be independently and identically distributed as $N(0,\sigma^2)$.

$$X_{ijk} = \beta_0 + \beta_1 TRT + \beta_2 PRD + \beta_{12}(TRT)(PRD) + s_{ij} + \varepsilon_{ijk}$$

The inclusion of the random subject effects term, the random intercept, would capture the within-subject correlation, called the intra-class correlation, $\sigma_s^2/(\sigma_s^2+\sigma^2)$ – as seen previously.

From the definitions of TRT and PRD, as compared to the notation of the previous method/lecture (A1, A2, B1, and B2), we have:

$$E(A1) = \beta_0 + \beta_1 + \beta_2 + \beta_{12}$$
$$E(A2) = \beta_0 + \beta_1 - \beta_2 - \beta_{12}$$
$$E(B1) = \beta_0 - \beta_1 + \beta_2 - \beta_{12}$$
$$E(B2) = \beta_0 - \beta_1 - \beta_2 + \beta_{12}$$

$$E(A1) = \beta_0 + \beta_1 + \beta_2 + \beta_{12}$$
$$E(A2) = \beta_0 + \beta_1 - \beta_2 - \beta_{12}$$
$$E(B1) = \beta_0 - \beta_1 + \beta_2 - \beta_{12}$$
$$E(B2) = \beta_0 - \beta_1 - \beta_2 + \beta_{12}$$

$$E(X1 = A1 - B2) = 2(\beta_1 + \beta_2)$$
$$E(X2 = A2 - B1) = 2(\beta_1 - \beta_2)$$
$$\frac{E(X1) + E(X2)}{2} = 2\beta_1$$
$$\frac{E(X1) - E(X2)}{2} = 2\beta_2$$

TREATMENT EFFECTS

X1 is normally distributed as $N(a+\beta,\sigma^2)$
X2 is normally distributed as $N(a-\beta,\sigma^2)$

$$\bar{x}_1 \text{ is } distributed \text{ as normal } N(\alpha+\beta, \frac{\sigma^2}{n}),$$

$$\bar{x}_2 \text{ is } distributed \text{ as normal } N(\alpha-\beta, \frac{\sigma^2}{n})$$

$$\mathbf{a} = \frac{\bar{\mathbf{x}}_1 + \bar{\mathbf{x}}_2}{\mathbf{2}}$$

$$\mathbf{a \text{ is } distributed \text{ as normal } N(\alpha, \frac{\sigma^2}{2n})}$$

ORDER EFFECTS

X1 is normally distributed as $N(\alpha+\beta,\sigma^2)$
X2 is normally distributed as $N(\alpha-\beta,\sigma^2)$

$$\bar{x}_1 \text{ is } distributed \text{ as normal } N(\alpha+\beta, \frac{\sigma^2}{n}),$$

$$\bar{x}_2 \text{ is } distributed \text{ as normal } N(\alpha-\beta, \frac{\sigma^2}{n})$$

$$\mathbf{b} = \frac{\bar{x}_1 - \bar{x}_2}{2}$$

$$\mathbf{b \text{ is } distributed \text{ as normal } N(\beta, \frac{\sigma^2}{2n})}$$

As far as Treatment Effect and Period effect are concerned, we get the same results using Simple Method or LME model

If we compare these results to those in the previous method/lecture, the results on the Treatment Effects and Period Effects are the same in both methods.

The question is what is new with LME Model and, especially, what does the interaction term, $\beta_{12}(TRT)(PRD)$, do?

$$E(A1) = \beta_0 + \beta_1 + \beta_2 + \beta_{12}$$
$$E(A2) = \beta_0 + \beta_1 - \beta_2 - \beta_{12}$$
$$E(B1) = \beta_0 - \beta_1 + \beta_2 - \beta_{12}$$
$$E(B2) = \beta_0 - \beta_1 - \beta_2 + \beta_{12}$$

$$E(A1 - A2) = 2(\beta_2 + \beta_{12})$$

The difference between A1 (Treatment A in Period 1) and A2 (Treatment A in Period 2) is more than just Period effects; the interaction term, represents the carried-over effects The use of LME Model allows us to investigate carried-over effects.

DESIGNS WITH A BINARY RESPONSE

Two-period crossover designs are often used in clinical trials in order to improved sensitivity of the trial by eliminating individual patient effects. They have been popular in dairy husbandry studies, long-term agricultural experiments, bioavailability and bioequivalence studies, nutrition experiments, arthritic and periodontal studies, and educational and psychological studies – where treatment effects are not permanent.

The response could be quantitative but quite often the response variable is binary , e.g. the response is whether or not relief from pain is obtained. We cover the case of binary responses more briefly.

THE DESIGN

Recall the following design; the only difference is that, in this case, the four outcomes A1, A2, B1, and B2 are binary – say 1 if positive response and 0 otherwise:

Group 1: Period #1 (Trt A; A1) – washout – Period #2 (Trt B; B2)
Group 2: Period #1 (Trt B; B1) – washout – Period #2 (Trt A; A2)

The washout periods are optional and the group sizes could be different (due to dropouts) – but not by much.

In general, let Y be the outcome or dependent variable taking on values 0 and 1, and:

π = Pr(Y=1)

Y is said to have the "Bernouilli distribution" (Binomial with n = 1). We have:

$$E(Y) = \pi$$

$$Var(Y) = \pi(1-\pi)$$

Studies would involve some independent variables (treatment, order, etc…)

LOGISTIC REGRESSION

Let π be the probability (also the mean of the Bernouilli distribution) and X a covariate (let consider only one X for simplicity). The common step in the regression modeling process is to relate π and X using the Logistic Regression Model – as follows.

$$\pi = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}}$$

$$1 - \pi = \frac{1}{1 + e^{\beta_0 + \beta_1 x}}$$

$$\frac{\pi}{1 - \pi} = e^{\beta_0 + \beta_1 x}$$

$$\log \frac{\pi}{1 - \pi} = \beta_0 + \beta_1 x$$

“Logistic Simple Linear Regression”

THE LOGISTIC MODELS

The Multiple Logistic Models for cross-over design are (J. J. Gart, Biometrika 1969):

$$\Pr(A1 = 1) = \frac{e^{\lambda_i + \alpha + \beta}}{1 + e^{\lambda_i + \alpha + \beta}}; \Pr(B2 = 1) = \frac{e^{\lambda_i - \alpha - \beta}}{1 + e^{\lambda_i + \alpha + \beta}}$$

$$\Pr(B1 = 1) = \frac{e^{\lambda_i - \alpha + \beta}}{1 + e^{\lambda_i - \alpha + \beta}}; \Pr(A2 = 1) = \frac{e^{\lambda_i + \alpha - \beta}}{1 + e^{\lambda_i + \alpha - \beta}}$$

In this models,
(1) λ‘s represent the subjects effects varying from subject to subject; could be many terms here.
(2) α represents the new treatment effect, say PEITC supplementation, (α>0 if and only if PEITC is more effective) – our main interest - and
(3) β represents the period effect (β>0 if and only if a treatment from period 1 is more effective than from period 2).

In this modeling:

(1) We “code” binary covariates (Treatment and Order) as (+1/-1) instead of (0,1);

(2) All subject-specific covariates are lumped together with the Intercept.

We want to eliminate the subjects' effects in drawing inferences on treatment and order effects. However, we cannot simply do some subtractions like (X1 = A1 - B2 and X2 = A2 - B1). For a continuous outcome, the difference of two normal variables is distributed as "normal". But this is not true for the Bernouilli distribution.

A viable alternative is a “Conditional Analysis”, like the formation of the McNemar Chi-square test – used, for example, in the analysis of pair-matched case-control studies.

It was shown by Gart (1969) that optimum inferences about treatment and order effects, regarding subjects effects as nuisance, are based on those subjects with unlike responses in two periods; that are subjects whose pair of outcomes are either (0,1) or (1,0). This is similar to the argument leading to the McNemar Chi-square test.

1: Period 1 (Trt A; A1) – Period 2 (Trt B; B2)
2: Period 1 (Trt B; B1) – Period 2 (Trt A; A2)

The analysis will be conditioned on:

A1+B2 = 1, and

B1+A2 = 1

$$\Pr(A1 = 1) = \frac{e^{\lambda_i + \alpha + \beta}}{1 + e^{\lambda_i + \alpha + \beta}}; \Pr(B2 = 1) = \frac{e^{\lambda_i - \alpha - \beta}}{1 + e^{\lambda_i + \alpha + \beta}}$$

$$\Pr(B1 = 1) = \frac{e^{\lambda_i - \alpha + \beta}}{1 + e^{\lambda_i - \alpha + \beta}}; \Pr(A2 = 1) = \frac{e^{\lambda_i + \alpha - \beta}}{1 + e^{\lambda_i + \alpha - \beta}}$$

Recall The Bayes Theorem:

$$\Pr(A1 = 1 \mid A1 + B2 = 1) = \frac{\Pr(A1 = 1, B2 = 0)}{\Pr(A1 = 1, B2 = 0) + \Pr(A1 = 0, B2 = 1)}$$

$$= \frac{\Pr(A1 = 1)\Pr(B2 = 0)}{\Pr(A1 = 1)\Pr(B2 = 0) + \Pr(A1 = 0)\Pr(B2 = 1)}$$

$$\Pr(A1 = 1 \mid A1 + B2 = 1) = \frac{\Pr(A1 = 1, B2 = 0)}{\Pr(A1 = 1, B2 = 0) + \Pr(A1 = 0, B2 = 1)}$$

$$= \frac{\Pr(A1 = 1)\Pr(B2 = 0)}{\Pr(A1 = 1)\Pr(B2 = 0) + \Pr(A1 = 0)\Pr(B2 = 1)}$$

$$= \frac{1}{1 + e^{-2(\alpha+\beta)}}$$

$$\Pr(A2 = 1 \mid B1 + A2 = 1) = \frac{1}{1 + e^{-2(\alpha-\beta)}}$$

$$\Pr(A1 = 1 \mid A1 + B2 = 1) = \frac{1}{1 + e^{-2(\alpha+\beta)}}$$

$$\Pr(A2 = 1 \mid B1 + A2 = 1) = \frac{1}{1 + e^{-2(\alpha-\beta)}}$$

$$\Pr(B1 = 1 \mid B1 + A2 = 1) = 1 - \frac{1}{1 + e^{-2(\alpha-\beta)}}$$

$$\Pr(B1 = 1 \mid B1 + A2 = 1) = \frac{1}{1 + e^{-2(-\alpha+\beta)}}$$

$$\Pr(A1 = 1 \mid A1 + B2 = 1) = \frac{1}{1 + e^{-2(\alpha+\beta)}} = p1$$

$$\Pr(A2 = 1 \mid B1 + A2 = 1) = \frac{1}{1 + e^{-2(\alpha-\beta)}} = p2$$

	Treatments (A,B)	Treatments (B,A)
Oucome A=1	y_{a1}	y_{a2}
Oucome B=1	y_{b1}	y_{b2}
Total	n_1	n_2

Results: With n_1 and n_2 fixed, y_{a1} and y_{a2} are distributed as Binomials $B(n_1,p1)$ and $B(n_2,p2)$

Results: With n_1 and n_2 fixed, y_{a1} and y_{a2} are distributed as Binomials $B(n_1,p1)$ and $B(n_2,p2)$

(Conditional) Likelihood Function:

$$L = \binom{n_1}{y_{a1}} p1^{y_{a1}} (1-p1)^{y_{b1}} \binom{n_2}{y_{a2}} p2^{y_{a2}} (1-p2)^{y_{b2}}$$

$$p1 = \frac{1}{1 + e^{-2(\alpha+\beta)}}$$

$$p2 = \frac{1}{1 + e^{-2(\alpha-\beta)}}$$

$$L = \binom{n_1}{y_{a1}} p1^{y_{a1}} (1 - p1)^{y1} \binom{n_2}{y_{a2}} p2^{y_{a2}} (1 - p2)^{y2}$$

$$= \frac{\binom{n_1}{y_{a1}} \binom{n_2}{y_{a2}} \left[\exp\{-2y_{b1}(\alpha + \beta) + 2y_{b2}(\beta - \alpha)\} \right]}{\left[1 + e^{-2(\alpha+\beta)}\right]^{n_1} \left[1 + e^{2(\beta-\alpha)}\right]^{n_2}}$$

RESULTS:
Estimates & Standard Errors

$$\hat{\alpha} = a = \frac{1}{4} \ln \frac{y_{a1} y_{a2}}{y_{b1} y_{b2}}$$

$$\hat{\beta} = b = \frac{1}{4} \ln \frac{y_{a1} y_{b2}}{y_{a2} y_{b1}}$$

$$\mathrm{Var}(a) = \mathrm{Var}(b) = \frac{1}{16} \left[\frac{n_{ab}}{y_{a1} y_{b1}} + \frac{n_{ba}}{y_{a2} y_{b2}} \right]$$

Suggested Exercise:

We conducted a randomized, crossover trial to test whether 3,3'-diindolylmethane (DIM, a metabolite of I3C) excreted in the urine after consumption of raw Brassica vegetables with divergent glucobrassicin concentrations is a marker of I3C uptake from such foods. Twenty-five subjects were fed 50 g of either raw "Jade Cross" Brussels sprouts (high glucobrassicin concentration) or "Blue Dynasty" cabbage (low glucobrassicin concentration) once daily for 3 days. All urine was collected for 24 hours after vegetable consumption each day. After a washout period, subjects crossed over to the alternate vegetable. Data are in file “Brussels Sprouts”; use average of 3 days as our outcome. Estimate & Test for Treatment effects using both the t-test (hand calculation) and SAS program (handout).