BIA 654 Homework 9

- Be sure to include statistical software outputs to show your work.
- Submit each homework at the beginning of each class.

Recall the case study on "Direct Mail Sales," using Plackett-Burman (PB) design; see recent lecture slides numbered 11–18.

(a) Consider the Plackett-Burman design in Table shown on slides 14 and 15. Check that the design is indeed orthogonal for the main effects. That is, pick any two factors, say A and F, and check that the design includes 5 runs at each of the four factor-level combinations. (Just check this for yourself and no need to put your answer here.)

(b) Reanalyze the data shown on slide 15 by the Plackett-Burman design; Excel data file is uploaded. More specifically,

  i. Produce the normal probability plot of 19 main effects and see what factors seem to be significant.

  ii. Produce also the Pareto plot as shown on slide number 28.

  iii. Screen out the 19 factors by examining their respective 95% confidence intervals.

• Here, I provide a hint to iii. Notice that the response variable is a *proportion* (in percent!) of customers who have responded (accepted the offer) to a particular mail sale. The total number N of people in this case study is $N = 100,000$.

For each effect, $N/2$ people will see the + level, and $N/2$ people will see the − level. Then the "effect" is the difference,

$$p_+ - p_-$$

where p_+ is the proportion of positive responses among the $N/2$ subjects exposed to the + level, and p_- is the proportion of positive responses among the $N/2$ subjects exposed to the − level.

Now, what is the standard error of effect?

$$\begin{aligned}\text{standard error (effect)} &= \sqrt{\text{estimate of Var}(p_+ - p_-)} \\ &= \sqrt{\text{estimate of Var}(p_+) + \text{estimate of Var}(p_-)}.\end{aligned}$$

Notice that the true variance $\text{Var}(p_+)$ is given by

$$\text{Var}(p_+) = \frac{\pi_+(1-\pi_+)}{N/2},$$

where π_+ is the *true* proportion of successes for customers exposed to the + level. Note: The above variance formula is from the fact $\text{Var}(\text{Binomial}(n,p)) = np(1-p)$ and so

$$\text{Var(proportion)=Var(Biomonial}(n,p)/n) = np(1-p)/n^2 = p(1-p)/n, \quad \text{and} \quad n = N/2.$$

Similarly, we calculate $\text{Var}(p_-)$. Under the null hypothesis $H_0 : \pi_+ = \pi_-$, standard error (effect) is given by

$$\text{standard error (effect)} = \sqrt{\frac{p(1-p)}{N/2} + \frac{p(1-p)}{N/2}},$$

where p is the *overall* success proportion, pooled over all runs. That is, in the case study, $p = 1.298\%$ by noting there are 1,298 people placed an order (positive response) out of $100,000$ people. You can use $z_{0.05/2} = 1.96$ in this case due to large sample size. Note the confidence intervals formed this way can be used to reject the null $H_0 : \pi_+ = \pi_-$ if the interval does not contain zero.

• **Caution:** When you construct confidence intervals, be careful about the unit, e.g., if the response is in percent, then standard error should also be in percent unit.