Introductory Statistical Analysis Session 8 Task

ASSIGNED ON: Wednesday 3/27/2013
DUE Tuesday 4/2/2013 9pm
You are to do the QUESTIONS IN BLUE for this Task.

Odds, Odds Ratios (OR): From Frequency to Probability (Pr)

Frequency: With no assumptions about the coin's fairness, toss it 10 times, heads show up in 3 tosses.
Odds of getting a head is 3/(10-3)=3/7
Translate frequency into probability:
3 -> 3/10 = 30%; 7 -> 7/10 = 70%; 10 = 100%
Calculate odds using probability: 0.3/0.7=3/7
Suppose I have another coin, out of 10 tosses 5 heads show up. Odds of getting head from coin2= ?
Now we have 2 odds. We can divide one by the other to get an ODDS RATIO (OR).

Compare odds and OR to probability

coin 1: Pr of head = 0.3	odds of head = 0.428
coin 2: Pr of head =	odds of head =
Pr ratio (coin1/coin2) for head: 0.6	OR (coin1/coin2) for head: 0.857
Pr ratio (coin2/coin1) for head: YOUR TURN	OR(coin2/coin1) for head: YOUR TURN

*UNLESS the coin is fair, odds and ORs are always numerically more extreme than Pr's and PR ratios

Between odds, OR, Pr and frequency

Can calculate Pr, odds and OR from frequency
Can calculate odds and OR if know Pr
CANNOT calculate frequency (actual data points) from only knowing Pr, odds or OR (summary statistics).
CAN calculate frequency from only knowing Pr, odds or OR IF KNOW RELEVANT SAMPLE SIZE.

Example: if only know coin 1 Pr of head = 0.3, cannot know # heads actually being tossed; If know coin 1 Pr of head = 0.3 AND there were a total of 10 tosses, can know # heads being tossed; If know coin 1 Pr of head = 0.3 AND 3 heads were tossed, can know total # coin tosses.

YOUR TURN: apply the same logic to odds—

Say you want to compare visual acuity between vitamin A-treated ADHD children and placebo (i.e. no vitamin A treatment). You measure your subjects' visual acuity using the Snellen eye chart by asking each subject 50 different letters on the chart. You and your statistician decide to MAKE THIS RATHER CONTINUOUS VARIABLE (number of letters correctly identified) INTO A DICHOTOMOUS VARIABLE, so that you could calculate odds, odds ratios, and model the outcome using logistic regression. You set the cut-off for "HIGH visual acuity" to be "correctly identifying at least 80% of the letters", and "LOW visual acuity" if a subject incorrectly identify MORE than 10 letters. Your statistician uses the number 1 for high visual acuity and 0 for low.

1. Write down either in mathematical expression or articulate in words how you calculate which OR (there are 2 possible ORs: OR as the result of using treatment group as denominator and OR as the result of using placebo group as denominator).

2. Say there are 100 subjects in treatment group and 200 subjects in placebo group. 80 of the subjects in treatment group had HIGH visual acuity after treatment. What is the odds that a subject in this group has LOW visual acuity? What is the probability that a subject in this group has LOW visual acuity?
3. Given that 80 out of 100 treated subjects had high visual acuity and 50 out of 200 placebo subjects had high visual acuity. What is the odds ratio (OR) calculated by dividing the treatment group's odds of high visual acuity by the placebo group's odds of high visual acuity?
4. If you know all the probability, odds, and odds ratios but NOTHING about the total subjects in each group and how many subjects in each group had high or low visual acuity, can you back-calculate to get the counts?

Interpreting OR

OR is a ratio.
OR(rain today / rain tomorrow)=1 means the odds that it will rain today and the odds that it will rain tomorrow are equal (both days could have 10%, 79%, etc chances of raining)
OR(rain today / rain tomorrow)=3 means the odds that it will rain today is 3 times the odds it will rain tomorrow
OR(rain today / rain tomorrow)=0 means the odds today will rain is either 0 or 100%
If OR(rain today / rain tomorrow)=3, OR(rain tomorrow / rain today)=1/3;
If OR(rain today / rain tomorrow)=0.2, OR(rain tomorrow / rain today)=1/0.2=5

Testing Ho in terms of OR

In terms of differences:
Ho states diff (group A vs B)=0;
In terms of ratios:
Ho states ratio(group A vs B)=1;
YOUR TURN: If I calculate ORs of low- divided by high-dose group's odds of positivity on Test SB (OR=1.3) and obtain a 95% confidence interval that is between 1.1 and 2.6:

1. Interpret this OR
2. Is there statistically significant difference between ___ from low-dose group and ___ from high-dose group?

OR calculation link: http://statpages.org/ctab2x2.html
*DO NOT FORGET odds are almost always more extreme than corresponding probabilities.

How to turn in this assignment:
E-mail me (li.xie@nemours.org) your file (any format) and name it:
LastnameFirstinitial_session8hw_mmddyyyy
For example, my name is Li Xie, so I would name my file "xiel_session8hw_04012013"

For any questions regarding either the concepts or the programming, please feel free to contact me.