

Statistical Reasoning in Public Health HW#4

Biostatistics 611, 2009

(Please note: Not only is it allowed, but I highly encourage you to use Stata for questions 1 and 2)

1. A 2003 article in the Journal of the American Medical Association¹ documents the results of a randomized clinical trial designed to evaluate whether the influenza vaccine is effective in reducing the occurrence of acute otitis media (AOM) in young children. Acute otitis media is an infection that causes inflammation of the middle ear canal. In the study, children were randomized to receive either the influenza vaccine or a placebo. (randomization was done in a 2 to 1 ratio, meaning that two times as many children were randomized to the vaccine treatment as were randomized to the placebo group). The children were followed for one year after randomization, and monitored for AOM during this period. 262 children were randomized to the vaccine group, and 150 of these children experienced at least one incident of AOM during the follow-up period. 134 children were randomized to the placebo group, and 83 of these children experience at least one incident of AOM during the follow-up period.
(18 points total: 3 points for each of a – f)
 - (a) Compute a 95% confidence interval for the proportion of children experiencing at least one incident of AOM during the follow-up period in each of the randomization groups. How do these 95% CI's compare? (similar range of values? Overlap?)
 - (b) Compute a 95% confidence interval for the difference in proportions of children experiencing AOM over the follow-up period. (you may choose the direction of comparison – ie: placebo to vaccine or vice-versa) . Interpret the confidence interval in a sentence.
 - (c) State the null and alternative hypotheses associated with testing for an association between the influenza vaccine and AOM. What is the p-value for testing these hypotheses? Is this p-value consistent with your result in part b?
 - (d) Give an estimate of the relative risk of AOM (in the follow-up period) for children in the vaccine group compared to those in placebo. Interpret this estimate.
 - (e) Give an estimate of the odds ratio of AOM (in the follow-up period) for children in the vaccine group compared to those in placebo. How does it compare in value to the estimate of relative risk from part d?
 - (f) Is this a randomized study? What does this suggest when translating the statistical result from part (b) into a substantive/scientific conclusion? *(Note: this is an extension question, we have not covered this in detail, I just want you to think about it)*

¹ based on data from:

Hoberman, A. et al. Effectiveness of Inactivated Influenza Vaccine in Preventing Acute Otitis Media in Young Children: A Randomized Controlled Trial (2003). *Journal of the American Medical Association*, 290, 1608-1616.

2. A study was done to investigate whether there is a relationship between survival of patients with coronary heart disease and pet ownership. A representative sample of 101 patients with CHD was taken. Each of these patients was classified as having a pet or not and by whether they survived one year following their first heart attack. Of 52 pet owners, 50 survived. Of 49 non-pet owners, 28 survived.

Suppose you were interested in doing a statistical analysis of these study results. Answering the follow questions to help you with this goal!

(21 points total: 3 points for each of a – g)

- (a) Using Stata, compute exact 95% confidence intervals for the proportion of patients surviving at least one year beyond the first heart attack in each of the groups (pet/no pet). How do these 95% CI's compare? (similar range of values? Overlap?)
- (b) By hand, use the CLT based approximation ($\hat{p} \pm 2se(\hat{p})$) to estimate the 95% confidence interval for the proportion of patients surviving at least one year beyond the first heart attack in each of the groups (pet/no pet). How do these two estimated CIs compare to the CIs computed by Stata in part (a)? Can you explain any large discrepancies?
- (c) Compute a 95% confidence interval for the difference in proportions of patients surviving at least one year beyond their first heart attack between the two groups. (you may use “large sample methods”, ie: the estimate given by the “csi” command in STATA, even if not fully appropriate)
- (d) State the null and alternative hypotheses associated with testing for an association between pet-ownership and one-year survival following a heart attack. What is the p-value for testing these hypotheses? Is this p-value consistent with your result in part (c)?
- (e) Give an estimate of the relative risk of survival (in the one year follow-up period) for subjects with pets compared to those without pets. Interpret this estimate.
- (f) Give an estimate of the odds of survival (in the one year follow-up period) for subjects with pets compared to those without pets. How does this compare in value to the relative risk estimate from e)?
- (g) Is this a randomized study? What does this suggest when translating the statistical result from part (b) into a substantive/scientific conclusion? *(Note: this is an extension question, we have not covered this in detail, I just want you to think about it)*

3. A pilot study was designed to evaluate the potential efficacy of a program designed to reduce prison recidivism amongst inmates who have a documented long-term history of drug and/or alcohol problems. A sample of 11 prisoners was followed for up to 24 months after their most recent release from prison. Six of the inmates returned to prison at 3, 7, 9, 11, 14 and 21 months respectively. Five of the inmates had not returned to prison since the last time they were last contacted since which was at 4, 8, 16, 24, and 24 months respectively. Estimate the survival curve describing return to prison for this sample of 11 inmates. You may either present this estimate as a table or a graphic. What statistical method are you using to do this? What is the estimated median time to return to prison for this group of 11 inmates?
(6 points)