Homework 5

Public Health 241: Statistical Analysis of Categorical Data

YOUR NAME / YOUR STUDENT ID HERE

TODAY'S DATE

1. The table below summarizes data from a traditional case-control study of oral cancer in females and employment in the textile industry.

Table 1: Oral Cancer

years	cases	controls
Ten or more years in textile industry	16	8
Other work history	24	32

(a) Calculate a point estimate for the odds ratio of oral cancer associated with having been employed in the textile industry for 10 or more years.

(b) Calculate an estimate of the variance of the log of your point estimate from (a).

(c) Construct an approximate 90% confidence interval for the log odds ratio of oral cancer associated with having been employed in the textile industry for 10 or more years. Hint: Use the base R function `qnorm()` to compute the correct z-value

(d) Construct an approximate 90% confidence interval for the odds ratio of oral cancer associated with having been employed in the textile industry for 10 or more years.

(e) Does the confidence interval provide evidence against the null hypothesis that employment history in the textile industry and the risk for oral cancer are independent of each other in the target population?

(f) Carry out the χ^2 test of independence. At a significance level of 10%, does the test accept or reject the null hypothesis of independence between employment history and the risk for oral cancer? Do these results agree with those based on the confidence interval calculation?

(g) Compute a point estimate for the odds ratio using the small-sample adjustment for obtaining a direct odds ratio estimate (rather than a point estimate on the log scale). Compare your point estimate to the one obtained in (a) and comment.

(h) Construct a 90% confidence interval for the odds ratio using the small-sample adjustment presented in class. Compare the results to those obtained above and comment.

(i) Obtain a