Homework 7

David Coomes

11/17/2019

Question 1

(a)

Table 1: Odds ratios and 95% confidence intervals for the effect of x-ray exposure adjusting for year of birth (using -10 for 1944).

	Estimate	95% CI Upper	95% CI Lower	P-value
YOB:1944	2.46	1.80	3.36	0.00
YOB:1948	2.11	1.71	2.60	0.00
YOB:1952	1.81	1.59	2.06	0.00
YOB:1956	1.55	1.38	1.75	0.00
YOB:1960	1.33	1.10	1.62	0.00
YOB:1964	1.14	0.85	1.53	0.38

(b)

(i)

Table 2: Odds ratios and 95% confidence intervals for the effect of x-ray exposure adjusting for year of birth (using -20 for 1944).

	Estimate	95% CI Upper	95% CI Lower	P-value
YOB:1944	2.46	1.80	3.36	0.00
YOB:1948	2.11	1.71	2.60	0.00
YOB:1952	1.81	1.59	2.06	0.00
YOB:1956	1.55	1.38	1.75	0.00
YOB:1960	1.33	1.10	1.62	0.00
YOB:1964	1.10	0.80	1.51	0.56

(ii)

The odds ratios are the same for both models except for the OR for those born in 1964. This model is not a reparamaterization - there is just one variable that is coded differently. Both models include all of the same variables. These models have a slightly different result because of the variable that is the baseline for the estimates.

(c)

Table 3: Odds ratios and 95% confidence intervals for the effect of x-ray exposure adjusting for year of birth using a quadratic effect modification model (using -10 for 1944).

	Estimate	95% CI Upper	95% CI Lower	P-value
YOB:1944	4.59	2.43	8.65	0.00
YOB:1948	2.51	1.93	3.27	0.00
YOB:1952	1.70	1.48	1.96	0.00
YOB:1956	1.43	1.24	1.64	0.00
YOB:1960	1.48	1.19	1.84	0.00
YOB:1964	1.90	1.11	3.22	0.02

(d)

Table 4: Odds ratios and 95% confidence intervals for the effect of x-ray exposure adjusting for year of birth using a quadratic effect modification model (using -20 for 1944).

	Estimate	95% CI Upper	95% CI Lower	P-value
YOB:1944	4.59	2.43	8.65	0.00
YOB:1948	2.51	1.93	3.27	0.00
YOB:1952	1.70	1.48	1.96	0.00
YOB:1956	1.43	1.24	1.64	0.00
YOB:1960	1.48	1.19	1.84	0.00
YOB:1964	1.81	1.05	3.12	0.03

These two quadratic effect modification models are very similar - they are only different for the OR in the final year (YOB 1966). This is the same as the two models that adjusted for effect modification using a non-quadratic equation - the differences are because of the chosen reference year.

2

(a) One way that the authors adjusted for confounders was in their study design. They limited the study to participants without diseases that could be related to alcohol and tobacco use, including diabetes and multiple cancers. As for adjusting using statistical methods, the authors are not very clear on the method used. They state their adjustment variables (in the results, not the methods) but not the way that they include them in the model. I would assume that they are all binary variables, mostly because some of the categories in Table IV have relatively few individuals (as low as 11), so this would not allow for a large number of confounders.

(b) I think that the presentation of the results was relatively clear, although there are a lot of them and I would have liked to have the main analysis emphasized a little more. It is also a little confusing that they changed their reference group from "Never drinkers" to "Current drinkers". I think that this makes the interpretation of the results a little more difficult to compare accross their analyses.

(c) I don't think that the authors can interpret as strong of a conclusion from this study as they did. For starters, there are a lot of confounding variables that are likely not accounted for in this study. It is very difficult to say that the cases and controls are equal except for the fact that the controls stopped drinking - there are too many differnet aspects that go along with drinking habits. Second, I wonder about the impact of mortality, especially for heavy drinkers, and what this may have done for the results of this study. What other diseases are associated with drinking and how would this impact the chances of someone dying before they get esophageal cancer. Finally, I wonder about the fact that such a large proportion of the study subjects were current drinkers and how this impacts the interpretation of the results.

Appendix

```
knitr::opts_chunk$set(warning = FALSE, message = FALSE)

library(knitr)
# install.packages('devtools', dependencies = TRUE, repos =
# 'http://cran.us.r-project.org')
library(devtools)
# install_github('bmckuw/UWbe536')
library(UWbe536)
library(kableExtra)
library(formatR)


link = "https://github.com/dmccoomes/Biostats_536/raw/master/Homework%207/oxford.rds"
ox <- readRDS(url(link))
head(ox)
summary(ox)
str(ox)

oxford <- ox[rep(1:nrow(ox), ox[, "freq"]), -4]
head(oxford)

with(oxford, table(case, xray))


library(dplyr)
oxford <- oxford %>% mutate(xraynum = as.numeric(xray) - 1, yob2 = yob *
    yob)


linintmod <- glm(case ~ yobf + xray + xraynum:yob, data = oxford,
    family = binomial)
coef(summary(linintmod))


est_1944 <- lincom(linintmod, c("xrayxray + (-10)*xraynum:yob==0"))
est_1948 <- lincom(linintmod, c("xrayxray + (-6)*xraynum:yob ==0"))
est_1952 <- lincom(linintmod, c("xrayxray + (-2)*xraynum:yob ==0"))
est_1956 <- lincom(linintmod, c("xrayxray + 2*xraynum:yob ==0"))
est_1960 <- lincom(linintmod, c("xrayxray + 6*xraynum:yob ==0"))
est_1964 <- lincom(linintmod, c("xrayxray + 10*xraynum:yob ==0"))



table_1 <- round(rbind(est_1944, est_1948, est_1952, est_1956,
    est_1960, est_1964), 2)
row.names(table_1) <- c("YOB:1944", "YOB:1948", "YOB:1952", "YOB:1956",
    "YOB:1960", "YOB:1964")


kable(table_1, col.names = c("Estimate", "95% CI Upper", "95% CI Lower",
    "P-value")) %>% kable_styling(full_width = F, position = "center")
```

```r
oxford$yob_2 <- oxford$yob - 10
summary(oxford$yob_2)


linintmod_2 <- glm(case ~ yobf + xray + xraynum:yob_2, data = oxford,
    family = binomial)
coef(summary(linintmod_2))


est2_1944 <- lincom(linintmod_2, c("xrayxray + (-20)*xraynum:yob_2==0"))
est2_1948 <- lincom(linintmod_2, c("xrayxray + (-16)*xraynum:yob_2 ==0"))
est2_1952 <- lincom(linintmod_2, c("xrayxray + (-12)*xraynum:yob_2 ==0"))
est2_1956 <- lincom(linintmod_2, c("xrayxray + (-8)*xraynum:yob_2 ==0"))
est2_1960 <- lincom(linintmod_2, c("xrayxray + (-4)*xraynum:yob_2 ==0"))
est2_1964 <- lincom(linintmod_2, c("xrayxray + xraynum:yob_2 ==0"))


table_2 <- round(rbind(est2_1944, est2_1948, est2_1952, est2_1956,
    est2_1960, est2_1964), 2)
row.names(table_2) <- c("YOB:1944", "YOB:1948", "YOB:1952", "YOB:1956",
    "YOB:1960", "YOB:1964")


kable(table_2, col.names = c("Estimate", "95% CI Upper", "95% CI Lower",
    "P-value")) %>% kable_styling(full_width = F, position = "center")


quadintmod <- glm(case ~ yobf + xray + xraynum:yob + xraynum:yob2,
    data = oxford, family = binomial)


est_1944 <- lincom(quadintmod, c("xrayxray + (-10)*xraynum:yob + 100*xraynum:yob2==0"))
est_1948 <- lincom(quadintmod, c("xrayxray + (-6)*xraynum:yob + 36*xraynum:yob2==0"))
est_1952 <- lincom(quadintmod, c("xrayxray + (-2)*xraynum:yob + 4*xraynum:yob2==0"))
est_1956 <- lincom(quadintmod, c("xrayxray + (2)*xraynum:yob + 4*xraynum:yob2 ==0"))
est_1960 <- lincom(quadintmod, c("xrayxray + (6)*xraynum:yob + 36*xraynum:yob2 ==0"))
est_1964 <- lincom(quadintmod, c("xrayxray + (10)*xraynum:yob + 100*xraynum:yob2 ==0"))


table_2 <- round(rbind(est_1944, est_1948, est_1952, est_1956,
    est_1960, est_1964), 2)
row.names(table_2) <- c("YOB:1944", "YOB:1948", "YOB:1952", "YOB:1956",
    "YOB:1960", "YOB:1964")


kable(table_2, col.names = c("Estimate", "95% CI Upper", "95% CI Lower",
    "P-value")) %>% kable_styling(full_width = F, position = "center")


quadintmod_2 <- glm(case ~ yobf + xray + xraynum:yob_2 + xraynum:yob2,
    data = oxford, family = binomial)
```

```r
est_1944 <- lincom(quadintmod_2, c("xrayxray + (-20)*xraynum:yob_2 + 100*xraynum:yob2==0"))
est_1948 <- lincom(quadintmod_2, c("xrayxray + (-16)*xraynum:yob_2 + 36*xraynum:yob2==0"))
est_1952 <- lincom(quadintmod_2, c("xrayxray + (-12)*xraynum:yob_2 + 4*xraynum:yob2==0"))
est_1956 <- lincom(quadintmod_2, c("xrayxray + (-8)*xraynum:yob_2 + 4*xraynum:yob2 ==0"))
est_1960 <- lincom(quadintmod_2, c("xrayxray + (-4)*xraynum:yob_2 + 36*xraynum:yob2 ==0"))
est_1964 <- lincom(quadintmod_2, c("xrayxray + xraynum:yob_2 + 100*xraynum:yob2 ==0"))


table_2 <- round(rbind(est_1944, est_1948, est_1952, est_1956,
    est_1960, est_1964), 2)
row.names(table_2) <- c("YOB:1944", "YOB:1948", "YOB:1952", "YOB:1956",
    "YOB:1960", "YOB:1964")


kable(table_2, col.names = c("Estimate", "95% CI Upper", "95% CI Lower",
    "P-value")) %>% kable_styling(full_width = F, position = "center")
```