

STATS402 HW3 Q6-10

Elbert Liu

2024-12-06

Setup

```
library(ggplot2)
library(finalfit)
```

```
## Warning: package 'finalfit' was built under R version 4.4.2
```

```
library(nnet)
```

```
## Warning: package 'nnet' was built under R version 4.4.2
```

```
df <- read.csv("C:\\Users\\Elbert\\OneDrive - UCLA IT Services\\402\\hw3\\liver23.csv")
df <- df[complete.cases(df),]
```

```
df <- df[(df$hist_hypertens_don.x != "U") & (df$coronary_angio_don.x != "U"),]
```

```
df$hist_hypertens_don.x <- factor(df$hist_hypertens_don.x)
df$coronary_angio_don.x <- factor(df$coronary_angio_don.x)
```

```
m1 <- glm(tx_fail ~ hgt_cm_don_calc.x + bmi_don_calc.x + coronary_angio_don.x + hist_hypertens_don.x, data = df)
```

I will be doing part 6 ~ 10 of Question 1. 6)

```
explanatory <- c("hgt_cm_don_calc.x", "bmi_don_calc.x", "coronary_angio_don.x", "hist_hypertens_don.x")
outcome <- "tx_fail"
df %>%
  or_plot(outcome, explanatory)
```

```
## Waiting for profiling to be done...
## Waiting for profiling to be done...
## Waiting for profiling to be done...
```

```
## Warning: Removed 2 rows containing missing values or values outside the scale range
## ('geom_errorbarh()').
```

tx_fail: OR (95% CI, p-value)

hgt_cm_don[1618,208.3] 0.97 (0.97–0.98, p<0.001)

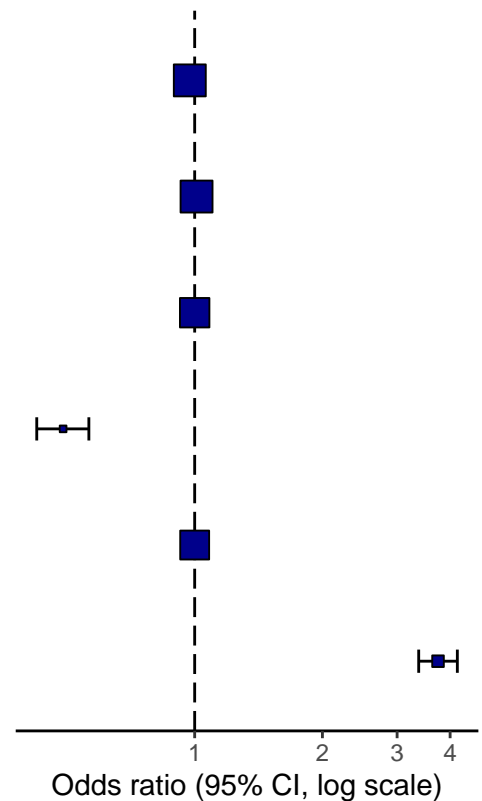
bmi_don_calc[15.1,66.4] 1.01 (1.00–1.02, p=0.022)

coronary_angio_don.x N

0.49 (0.42–0.56, p<0.001)

hist_hypertens_don.x N

3.75 (3.37–4.16, p<0.001)



7) The odds table shows that bmi and height does not cause the odds of transplant success to deviate from 1 that much. On the other hand, if the donor died from coronary heart disease, we are 95% sure that the odds of successful transplant to failure is between .42 and .56. If the donor had history of hypertension, we're 95% sure that the odds of successful transplant to failure is between 3.37 and 4.16.

8)

```
null_dev <- 12358
res_dev <- 11452
diff <- null_dev - res_dev
dof <- 4 #we have 4 degrees of freedom because we have 4 explanatory variables
print(pchisq(diff,dof))
```

```
## [1] 1
```

This basically means that the p value is 0 and our model is statistically significant in improving the predictability of transplant success.

9)

```
pseudo_rsqr <- 1 - (res_dev / null_dev)
print(pseudo_rsqr)
```

```
## [1] 0.07331283
```

10)

```
m2 <- multinom(tx_fail ~ hgt_cm_don_calc.x + bmi_don_calc.x + coronary_angio_don.x + hist_hypertens_don
```

```
## # weights:  6 (5 variable)
## initial  value 8631.761840
## iter   10 value 5727.139854
## final   value 5726.040448
## converged
```

```
p <- predict(m2,df)
tab <- table(p,df$tx_fail)
print(tab)
```

```
##
## p      0      1
##  0 9852 2253
##  1  148  200
```