step1单因素分析

Cindy

2025-02-26

# 加载必要的库  
library(tidyverse)

## ── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0 ──  
## ✔ dplyr 1.1.4 ✔ readr 2.1.5  
## ✔ forcats 1.0.0 ✔ stringr 1.5.1  
## ✔ ggplot2 3.5.1 ✔ tibble 3.2.1  
## ✔ lubridate 1.9.4 ✔ tidyr 1.3.1  
## ✔ purrr 1.0.2   
## ── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
## ✖ dplyr::filter() masks stats::filter()  
## ✖ dplyr::lag() masks stats::lag()  
## ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors

library(readr)  
library(dplyr)  
# 检测文件编码  
encoding\_info <- guess\_encoding("C://Users//Intern1//Desktop//数据2.csv")  
print(encoding\_info)

## # A tibble: 4 × 2  
## encoding confidence  
## <chr> <dbl>  
## 1 GB18030 0.93  
## 2 EUC-KR 0.33  
## 3 EUC-JP 0.24  
## 4 Big5 0.24

# 根据检测结果读取文件  
data <- read.csv("C://Users//Intern1//Desktop//数据2.csv",   
 stringsAsFactors = TRUE,   
 fileEncoding = encoding\_info$encoding[which.max(encoding\_info$confidence)])  
head(data)

## 年龄 高龄产妇 人工受精 孕次 产次 孕次分层 生育过 流产史 胎儿生长受限史  
## 1 27 0 0 1 0 0 0 0 0  
## 2 34 0 0 1 0 0 0 0 0  
## 3 26 0 0 1 0 0 0 0 0  
## 4 34 0 0 7 1 1 1 1 0  
## 5 24 0 0 1 0 0 0 0 0  
## 6 28 0 0 3 0 1 0 1 0  
## 高血压疾病 妊娠期肝内胆汁淤积症 糖尿病 甲状腺疾病 前置胎盘 易栓症  
## 1 0 1 0 0 0 0  
## 2 0 0 0 0 0 0  
## 3 0 0 0 0 0 0  
## 4 0 0 0 0 0 0  
## 5 0 0 0 0 0 0  
## 6 0 0 1 1 0 0  
## 抗心磷脂抗体综合征 脐血流异常 羊水少 FGR其他相关合并症  
## 1 0 0 0 0  
## 2 0 0 0 0  
## 3 0 0 1 0  
## 4 0 0 1 0  
## 5 0 0 1 0  
## 6 0 0 1 0  
## 既往小于10周难免流产的次数 既往小于10周难免流产的次数.1  
## 1 0 0  
## 2 0 0  
## 3 0 0  
## 4 0 0  
## 5 0 0  
## 6 0 0  
## 既往小于10周难免流产的次数.2 既往有无大于10周流产或死胎 身高  
## 1 1 0 1.64  
## 2 1 0 1.63  
## 3 1 0 1.58  
## 4 1 0 1.58  
## 5 1 0 1.70  
## 6 1 0 1.60  
## 本次妊娠孕前体重Kg 分娩前体重Kg 孕期体重增加Kg 孕前体重指数 孕前体重指数.1  
## 1 42.0 53.0 11.0 15.62 0  
## 2 53.0 65.5 12.5 19.95 1  
## 3 48.0 58.0 10.0 19.23 1  
## 4 60.0 71.0 11.0 24.03 2  
## 5 50.5 60.0 9.5 17.47 0  
## 6 50.0 62.5 12.5 19.53 1  
## 孕期体重增加是否符合相应BMI值 首次孕周 重度FGR heparinuse 肝素使用时间  
## 1 0 31.00000 1 1 28  
## 2 1 33.00000 1 1 21  
## 3 0 32.14286 1 1 35  
## 4 1 31.57143 0 1 49  
## 5 0 34.00000 1 1 17  
## 6 1 28.28571 1 1 40  
## 住院治疗 络安命 维生素C 硫酸镁 地塞米松 性别 出生体重g 出生1分钟评分  
## 1 1 1 0 1 0 2 2300 9  
## 2 1 1 1 1 1 2 1800 6  
## 3 1 1 1 1 1 2 2010 10  
## 4 1 1 0 1 0 2 2500 10  
## 5 1 1 1 0 0 2 1700 10  
## 6 1 1 1 1 1 2 2280 10  
## 出生5分钟评分 入住NICU 分娩孕周 新生儿评分 胎盘早剥 重度子痫前期 HELLP  
## 1 7 1 2 1 0 0 0  
## 2 8 1 3 1 0 0 0  
## 3 10 0 3 0 0 0 0  
## 4 10 0 3 0 0 0 0  
## 5 10 0 2 0 0 0 0  
## 6 10 0 2 0 0 0 0  
## fgroutcome 分娩时是否为重度FGR 分娩孕周.1 由首次及最后诊断  
## 1 0 0 36 3  
## 2 1 1 37 1  
## 3 1 1 37 1  
## 4 1 0 37.71428571 0  
## 5 1 1 36.42857143 1  
## 6 1 0 36.28571429 3

# 将分类变量转换为因子  
data$高龄产妇 <- as.factor(data$高龄产妇)  
data$人工受精 <- as.factor(data$人工受精)  
data$孕次分层 <- as.factor(data$孕次分层)  
data$生育过 <- as.factor(data$生育过)  
data$流产史 <- as.factor(data$流产史)  
data$胎儿生长受限史 <- as.factor(data$胎儿生长受限史)  
data$高血压疾病 <- as.factor(data$高血压疾病)  
data$妊娠期肝内胆汁淤积症 <- as.factor(data$妊娠期肝内胆汁淤积症)  
data$糖尿病 <- as.factor(data$糖尿病)  
data$甲状腺疾病 <- as.factor(data$甲状腺疾病)  
data$前置胎盘 <- as.factor(data$前置胎盘)  
data$易栓症 <- as.factor(data$易栓症)  
data$抗心磷脂抗体综合征 <- as.factor(data$抗心磷脂抗体综合征)  
data$脐血流异常 <- as.factor(data$脐血流异常)  
data$羊水少 <- as.factor(data$羊水少)  
data$FGR其他相关合并症 <- as.factor(data$FGR其他相关合并症)  
data$既小于10周难免流产的次数 <- as.factor(data$既往小于10周难免流产的次数)  
data$既往有无大于10周流产或死胎 <- as.factor(data$既往有无大于10周流产或死胎)  
data$fgroutcome <- as.factor(data$fgroutcome)  
data$heparinuse <- as.factor(data$heparinuse)  
data$住院治疗 <- as.factor(data$住院治疗)  
data$络安命 <- as.factor(data$络安命)  
data$维生素C <- as.factor(data$维生素C)  
data$硫酸镁 <- as.factor(data$硫酸镁)  
data$地塞米松 <- as.factor(data$地塞米松)  
data$heparinuse <- as.factor(data$heparinuse)  
# 设置因变量和自变量列表  
dependent\_var <- "fgroutcome"  
independent\_vars <- c("年龄", "高龄产妇", "人工受精", "孕次", "产次", "孕次分层", "生育过", "流产史", "胎儿生长受限史", "高血压疾病", "妊娠期肝内胆汁淤积症", "糖尿病", "甲状腺疾病", "前置胎盘", "易栓症", "抗心磷脂抗体综合征", "脐血流异常", "羊水少", "FGR其他相关合并症", "既往小于10周难免流产的次数", "既往有无大于10周流产或死胎", "身高", "本次妊娠孕前体重Kg", "分娩前体重Kg", "孕期体重增加Kg", "孕前体重指数", "孕期体重增加是否符合相应BMI值", "首次孕周", "肝素使用时间", "住院治疗", "络安命", "维生素C", "硫酸镁", "地塞米松","heparinuse")  
  
# 创建一个空的列表来存储结果  
results <- list()  
  
# 批量进行单因子逻辑回归分析  
for (var in independent\_vars) {  
 # 构建模型公式  
 formula <- as.formula(paste(dependent\_var, "~", var))  
   
 # 拟合逻辑回归模型  
 model <- glm(formula, data = data, family = binomial)  
   
 # 获取模型摘要  
 model\_summary <- summary(model)  
   
 # 提取系数和 p 值  
 coef\_data <- data.frame(  
 Variable = var,  
 Coefficient = model\_summary$coefficients[2, "Estimate"],  
 Std.Error = model\_summary$coefficients[2, "Std. Error"],  
 z.value = model\_summary$coefficients[2, "z value"],  
 p.value = model\_summary$coefficients[2, "Pr(>|z|)"]  
 )  
   
 # 将结果添加到列表中  
 results[[var]] <- coef\_data  
}  
  
# 合并所有结果为一个数据框  
results\_df <- do.call(rbind, results)  
  
# 打印结果  
print(results\_df)

## Variable Coefficient  
## 年龄 年龄 0.0049031600  
## 高龄产妇 高龄产妇 0.3953127040  
## 人工受精 人工受精 -0.4593652310  
## 孕次 孕次 0.1503782600  
## 产次 产次 0.0399154649  
## 孕次分层 孕次分层 0.5292309647  
## 生育过 生育过 0.0124225200  
## 流产史 流产史 0.1559948032  
## 胎儿生长受限史 胎儿生长受限史 0.9492476800  
## 高血压疾病 高血压疾病 0.5774506556  
## 妊娠期肝内胆汁淤积症 妊娠期肝内胆汁淤积症 0.0216850915  
## 糖尿病 糖尿病 -0.3230871229  
## 甲状腺疾病 甲状腺疾病 0.2231435476  
## 前置胎盘 前置胎盘 -0.5440040224  
## 易栓症 易栓症 -1.6184357714  
## 抗心磷脂抗体综合征 抗心磷脂抗体综合征 0.1800514068  
## 脐血流异常 脐血流异常 0.5880481370  
## 羊水少 羊水少 -0.1056125000  
## FGR其他相关合并症 FGR其他相关合并症 -0.2347043737  
## 既往小于10周难免流产的次数 既往小于10周难免流产的次数 -0.1430932263  
## 既往有无大于10周流产或死胎 既往有无大于10周流产或死胎 13.6762784349  
## 身高 身高 -0.0883111040  
## 本次妊娠孕前体重Kg 本次妊娠孕前体重Kg 0.0123476053  
## 分娩前体重Kg 分娩前体重Kg -0.0006376692  
## 孕期体重增加Kg 孕期体重增加Kg -0.0418993865  
## 孕前体重指数 孕前体重指数 0.0361953072  
## 孕期体重增加是否符合相应BMI值 孕期体重增加是否符合相应BMI值 -0.3129139126  
## 首次孕周 首次孕周 -0.1358635395  
## 肝素使用时间 肝素使用时间 -0.0322009190  
## 住院治疗 住院治疗 -0.5827115027  
## 络安命 络安命 -1.5819747157  
## 维生素C 维生素C -0.6008718987  
## 硫酸镁 硫酸镁 -0.8219089863  
## 地塞米松 地塞米松 0.1240687476  
## heparinuse heparinuse -1.3572444432  
## Std.Error z.value p.value  
## 年龄 2.419817e-02 0.20262522 8.394280e-01  
## 高龄产妇 3.275574e-01 1.20685019 2.274898e-01  
## 人工受精 4.095113e-01 -1.12174000 2.619730e-01  
## 孕次 9.134864e-02 1.64620149 9.972228e-02  
## 产次 1.869177e-01 0.21354562 8.309014e-01  
## 孕次分层 2.732501e-01 1.93680049 5.276974e-02  
## 生育过 2.367591e-01 0.05246904 9.581550e-01  
## 流产史 2.334163e-01 0.66831141 5.039348e-01  
## 胎儿生长受限史 1.035312e+00 0.91687135 3.592101e-01  
## 高血压疾病 6.212311e-01 0.92952632 3.526164e-01  
## 妊娠期肝内胆汁淤积症 4.526653e-01 0.04790536 9.617917e-01  
## 糖尿病 2.935493e-01 -1.10062322 2.710607e-01  
## 甲状腺疾病 4.481628e-01 0.49790730 6.185494e-01  
## 前置胎盘 4.713063e-01 -1.15424739 2.483988e-01  
## 易栓症 4.497213e-01 -3.59875277 3.197470e-04  
## 抗心磷脂抗体综合征 1.066776e+00 0.16878097 8.659689e-01  
## 脐血流异常 3.888476e-01 1.51228445 1.304615e-01  
## 羊水少 2.536613e-01 -0.41635246 6.771521e-01  
## FGR其他相关合并症 6.396760e-01 -0.36691136 7.136851e-01  
## 既往小于10周难免流产的次数 1.944182e-01 -0.73600720 4.617263e-01  
## 既往有无大于10周流产或死胎 5.500886e+02 0.02486196 9.801651e-01  
## 身高 2.211654e+00 -0.03992990 9.681490e-01  
## 本次妊娠孕前体重Kg 1.317348e-02 0.93730798 3.486002e-01  
## 分娩前体重Kg 1.160922e-02 -0.05492781 9.561960e-01  
## 孕期体重增加Kg 2.458992e-02 -1.70392522 8.839506e-02  
## 孕前体重指数 3.590368e-02 1.00812255 3.133956e-01  
## 孕期体重增加是否符合相应BMI值 2.310083e-01 -1.35455723 1.755587e-01  
## 首次孕周 6.365443e-02 -2.13439245 3.281067e-02  
## 肝素使用时间 7.040066e-03 -4.57395125 4.786112e-06  
## 住院治疗 2.906672e-01 -2.00473742 4.499113e-02  
## **络安命 3.211535e-01 -4.92591442 8.396674e-07**## 维生素C 2.295341e-01 -2.61778936 8.850142e-03  
## 硫酸镁 2.772592e-01 -2.96440687 3.032670e-03  
## 地塞米松 2.278959e-01 0.54440969 5.861596e-01  
## **heparinuse 2.337806e-01 -5.80563440 6.412274e-09**