

R 语言实现基于 Bliss 法的 LD_{50} 测定

shujuecn

更新: 2021 年 05 月 29 日

1 预设

1.1 宏包与字体

导入依赖宏包, 设置绘图中文字体。

```
library(tidyverse)
library(ggpmisc)
library(showtext)
# setwd("/your/files/path")
showtext_auto(enable = TRUE)
font_add('SimHei', regular = "/fontset/SimHei.ttf")
```

1.2 数据预处理

将所需数据整理, 构建数据框。

```
meta_data <- tibble(
  number = c(1:5),
  n = rep(8, 5),
  drug_rose = c(110.8, 147.7, 196.9, 262.5, 350.0),
  log_dose = log10(drug_rose),      # x
  lethal_rate = c(0, 0.25, 0.50, 0.625, 0.875),
  lethal_rate_p = qnorm(lethal_rate) + 5    # y
)
attach(meta_data)
meta_data

## # A tibble: 5 x 6
##   number      n drug_rose log_dose lethal_rate lethal_rate_p
```

##	<int>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
## 1	1	8	111.	2.04	0	-Inf
## 2	2	8	148.	2.17	0.25	4.33
## 3	3	8	197.	2.29	0.5	5
## 4	4	8	262.	2.42	0.625	5.32
## 5	5	8	350	2.54	0.875	6.15

2 统计分析

2.1 拟合第一条 LD-P 直线

选择死亡率 (lethal_rate) 不为 0 或 1 的组别, 以对数剂量 (log_dose) 为 X, 以死亡概率单位 (lethal_rate_p) 为 Y, 使用 lm 函数拟合第一条 LD-P 回归直线。

```
result_lm1 <- with(meta_data,
  lm(lethal_rate_p[2:5] ~ log_dose[2:5])
)
summary(result_lm1)
```

```
##
## Call:
## lm(formula = lethal_rate_p[2:5] ~ log_dose[2:5])
##
## Residuals:
##      1      2      3      4
## -0.004207  0.091096 -0.169536  0.082647
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -5.7329     1.2523  -4.578  0.0445 *
## log_dose[2:5]  4.6385     0.5304   8.745  0.0128 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1481 on 2 degrees of freedom
## Multiple R-squared:  0.9745, Adjusted R-squared:  0.9618
## F-statistic: 76.47 on 1 and 2 DF, p-value: 0.01283
```

2.1.1 建立直线回归方程

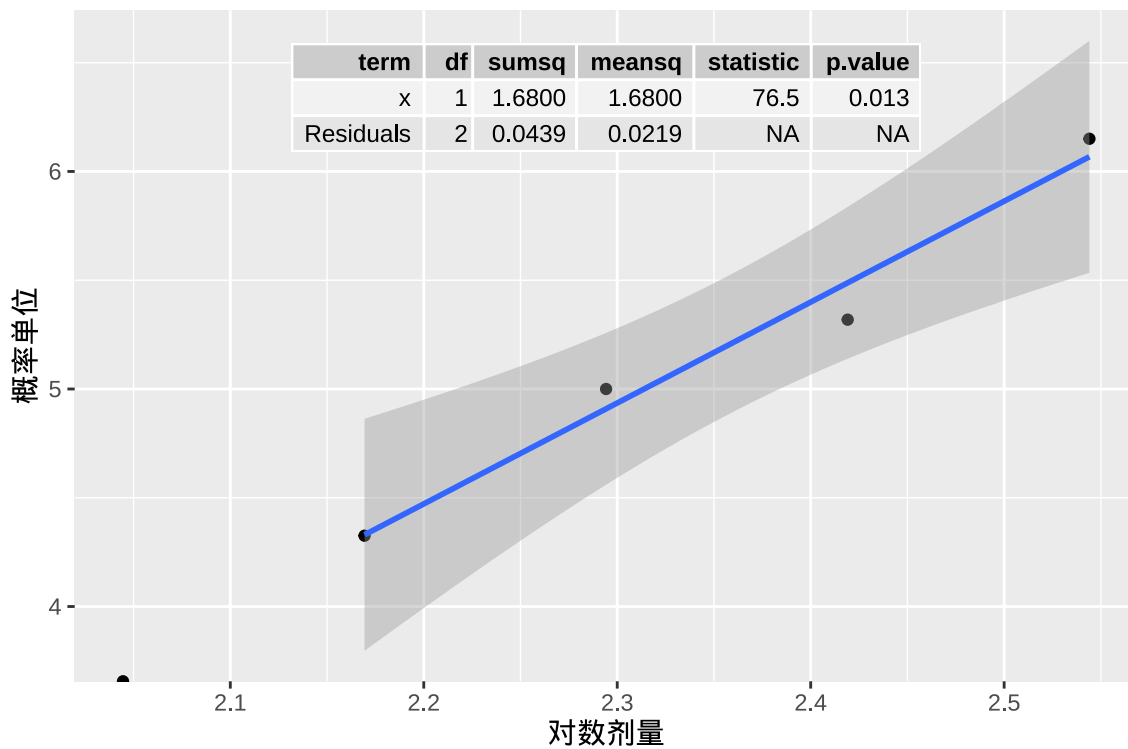
由结果可知，截距 a 为-5.7329，回归系数 b 为 4.6385，故第一条回归直线的方程式为：

$$\hat{Y} = -5.7329 + 4.6385X$$

2.1.2 可视化演示

绘制散点与回归直线图，如下所示：

```
meta_data %>%
  ggplot(aes(x = log_dose, y = lethal_rate_p)) +
  geom_point(aes(x = log_dose, y = lethal_rate_p)) +
  geom_smooth(method = 'lm', formula = y ~ x) +
  stat_fit_tb(tb.type = 'fit.anova') + # 方差分析表
  xlab(" 对数剂量") +
  ylab(" 概率单位")
```



2.2 校正 LD-P 回归直线

2.2.1 计算作业死亡概率单位

各组的作业死亡几率值 (y_r):

```

a <- coef(result_lm1)[1] # LD-P_1 的截距
b <- coef(result_lm1)[2] # LD-P_1 的回归系数
ye <- a + b * log_dose    # 各组的期望死亡几率值
z <- (ye-5) %>% dnorm()   # 正态分布概率
p <- (ye-5) %>% pnorm()   # 标准正态分布的密度函数
yr <- ye - p/z + lethal_rate/z # 作业死亡几率值
yr

```

```
## [1] 3.172035 4.325516 5.000252 5.312287 6.146720
```

2.2.2 计算权重系数

各组的权重系数 (w):

```

w <- (z**2) / (p * (1 - p))
w

```

```
## [1] 0.3532903 0.5397043 0.6347025 0.5834951 0.4157665
```

2.2.3 加权直线回归

以对数剂量 (log_dose) 为 X, 以作业死亡几率 (yr) 为 Y, 以权重系数 (w) 为各组权重, 使用 lm 函数拟合校正 LD-P 回归直线。

```

result_lm2 <- lm(yr ~ log_dose, weights = w)
summary(result_lm2)

```

```

##
## Call:
## lm(formula = yr ~ log_dose, weights = w)
##
## Weighted Residuals:
##      1      2      3      4      5
## -0.178520  0.133175  0.146608 -0.134465 -0.009017
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -7.5308      1.5579  -4.834  0.01688 *
## log_dose       5.3817      0.6749   7.974  0.00412 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

```
##
## Residual standard error: 0.1725 on 3 degrees of freedom
## Multiple R-squared:  0.9549, Adjusted R-squared:  0.9399
## F-statistic: 63.58 on 1 and 3 DF,  p-value: 0.004116
```

2.2.4 建立校正直线回归方程

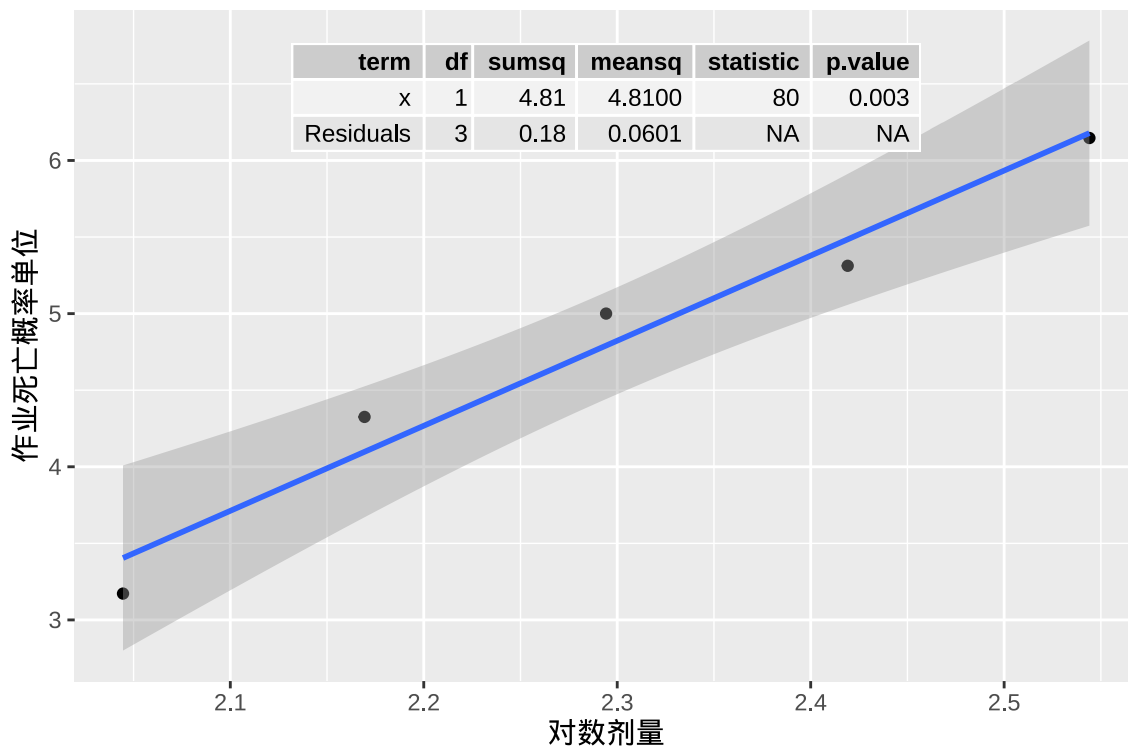
由结果可知，截距 a 为-7.5308，回归系数 b 为 5.3817，故第一条回归直线的方程式为：

$$\hat{Y} = -7.5308 + 5.3817X$$

2.2.5 可视化演示

绘制散点与回归直线图，如下所示：

```
meta_data %>%
  ggplot(aes(x = log_dose, y = yr)) +
  geom_point(aes(x = log_dose, y = yr)) +
  geom_smooth(method = 'lm', formula = y ~ x) +
  stat_fit_tb(tb.type = 'fit.anova') + # 方差分析表
  xlab(" 对数剂量") +
  ylab(" 作业死亡概率单位")
```



2.3 测定半数致死量

2.3.1 计算 LD_{50}

由回归直线方程式，计算半数致死量对数：

```
m <- (5 - coef(result_lm2)[1]) / coef(result_lm2)[2]
m
```

```
## (Intercept)
##      2.328393
```

$lgLD_{50}$ 取反对数，计算半数致死量：

```
10 ** m
```

```
## (Intercept)
##      213.0066
```

2.3.2 计算置信区间

计算 $lgLD_{50}$ 的标准误：

```
sum_n_w <- sum(n * w)
sum_n_w_x <- sum(n * w * log_dose)
sum_n_w_xx <- sum(n * w * log_dose**2)
mean_x <- sum_n_w_x / sum_n_w
sxx <- sum_n_w_xx - sum_n_w_x**2 / sum_n_w
se <- (sqrt(1 / sum_n_w + (m - mean_x)**2 / sxx)) / b
se
```

```
## (Intercept)
##  0.04856178
```

计算 95%CI:

```
10**(m + 1.96 * c(-1, 1) * se)
```

```
## [1] 171.0850 265.2003
```

3 结论

本药物的半数致死量为 213.0(mg/kg)，95% 置信区间为 [171.09, 265.20]。