

# 使用R进行中介分析

Alex / 2022-07-01 / [free\\_learner@163.com](mailto:free_learner@163.com) / [AlexBrain.cn](http://AlexBrain.cn)

更新于2023-09-27，主要是文字排版上的更新，内容基本保持不变。

测试了 `lavaan`、`mediation` 和 `psych` 包的中介分析（mediation analysis）用法，并比较一下结果差异。

## 一、原理和测试数据

关于中介分析的原理，请见参考资料1-2。[测试数据](#)来源见参考资料3。测试数据包含3个变量（如下图所示），`room_temp` 表示室温，`thirst` 表示口渴程度，`consume` 表示饮水量，这里要检验的是室温和饮水量的关系是否受到口渴程度中介。

<code>room_temp</code>	<code>thirst</code>	<code>consume</code>
69	2	3
70	2	4
69	1	2
70	3	2
69	1	1
70	2	3

## 二、使用不同R包进行分析

这里我测试了 `lavaan`、`mediation` 和 `psych` 包，也手工实现了一下。

```

## load test dataset
my_dat <- read.csv('mackinnon_2008_t10.1.csv')
## use lavaan package
library(lavaan)
set.seed(100)
my_mod <- "# a path
          thirst ~ 1 + a * room_temp
          # b path
          consume ~ 1 + b * thirst
          # c prime path
          consume ~ cp * room_temp
          # indirect and total effects
          ab := a * b
          total := cp + ab
"
my_sem.fit <- sem(my_mod, data = my_dat, se = "bootstrap", bootstrap = 1000)
summary(my_sem.fit)
lavaan_result <- parameterestimates(my_sem.fit, boot.ci.type = "perc", standardized = TRUE)
lavaan_result
## use mediation package
library(mediation)
set.seed(100)
my_mod.m <- lm(thirst ~ room_temp, data = my_dat)
my_mod.y <- lm(consume ~ room_temp + thirst, data=my_dat)
mediation_result <- mediation::mediate(my_mod.m, my_mod.y, treat='room_temp',
mediator='thirst', sims=1000, boot=TRUE, boot.ci.type = 'perc')
summary(mediation_result)
## use psych package
## the set.seed is of no use
library(psych)
set.seed(100)
psych_result <- psych::mediate(consume ~ room_temp + (thirst) , data=my_dat, n.iter=1000)
summary(psych_result)
## manual method
set.seed(100)
my_mod.y <- summary(lm(consume ~ room_temp + thirst, data = my_dat))
my_mod.m <- summary(lm(thirst ~ room_temp, data = my_dat))
a <- my_mod.m$coefficients[2,1]
a_se <- my_mod.m$coefficients[2,2]
b <- my_mod.y$coefficients[3,1]
b_se <- my_mod.y$coefficients[3,2]
## Sobel's Z-test
ab <- a * b
ab_se <- sqrt(b^2 * a_se^2 + a^2 * b_se^2)
Z <- ab / (ab_se)
p <- 2 * pnorm(-abs(Z))
## bootstrap

```

```

boot_ab <- NULL
N <- nrow(my_dat)
sample_idx <- 1:N
nboot=1000
boot_idx <- replicate(nboot, sample(sample_idx, size = N, replace = TRUE))
for (i in c(1:nboot)){
  boot_dat <- my_dat[boot_idx[,i],]
  boot_mod.y <- lm(consume ~ room_temp + thirst, data = boot_dat)
  boot_mod.m <- lm(thirst ~ room_temp, data = boot_dat)
  boot_ab <- c(boot_ab, boot_mod.y$coefficients[3] * boot_mod.m$coefficients[2])
}
CI <- quantile(boot_ab, probs = c(0.025, 0.975))
manual_result <- data.frame(ab=ab, ab_se=ab_se, Z=Z, p=p, lb=CI[1], ub=CI[2])
manual_result

```

### 三、不同 R 包的结果比较

如下表所示，我只关注了间接效应（indirect effect）的估计值、标准误、Z值、p值和95%置信区间。不同R包的输出的内容和结果稍有不同。值得注意的是 `psych` 包，尽管我设置了seed，每次跑的结果都不一样，不太清楚原因。这里我只比较了最简单的单个中介变量的模型结果，更复杂的模型结果如何，有空再测试。

Method	ab	SE	Z	p	95% CI
lavaan	0.26	0.13	2.10	0.035	(0.034, 0.53)
mediation	0.26	NA	NA	0.022	(0.052, 0.52)
psych	0.26	0.12	NA	NA	(0.040, 0.52)
manual	0.26	0.13	2.08	0.037	(0.037, 0.53)

### 四、参考资料

1. <https://advstats.psychstat.org/book/mediation/index.php>
2. MacKinnon, D. P., Fairchild, A. J., & Fritz, M. S. (2007). Mediation analysis. *Annual review of psychology*, 58, 593–614. <https://doi.org/10.1146/annurev.psych.58.110405.085542>
3. <https://nickmichalak.com/post/2019-02-13-testing-indirect-effects-mediation-in-r/testing-indirect-effects-mediation-in-r/>