Stat 3202 Lab 7

Jane Weissberg.11

2024-07-05

1a.

```
A <- c(25.6,23.2,26.5,25.8,24.7,23.8,25.3,24.1,24.2,26.2)
B <- c(29,29.3,31.1,29.1,29.5,29.6,27.6,27.9,28.5,30)

## Compute a 95% CI assuming variances are equal and unknown

nA <- length(A)
nB <- length(B)
alpha <- 0.05
xbar_a <- mean(A)
xbar_b <- mean(B)
s2a <- var(A)
s2b <- var(B)
s2pooled <- ((nA-1)*s2a+(nB-1)*s2b)/(nA+nB-2)
tstar <- qt(1-alpha/2, nA+nB-2)

LL <- (xbar_a-xbar_b)-tstar*sqrt(s2pooled/nA+s2pooled/nB)
UL <- (xbar_a-xbar_b)+tstar*sqrt(s2pooled/nA+s2pooled/nB)
round(c(LL, UL), 2)
```

```
## [1] -5.22 -3.22
```

```
library(DescTools)
t.test(A, B, paired=FALSE, var.equal=TRUE, conf.level=0.95)
```

```
##
##  Two Sample t-test
##
## data:  A and B
## t = -8.897, df = 18, p-value = 5.226e-08
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -5.216506 -3.223494
## sample estimates:
## mean of x mean of y
##     24.94     29.16
```

1b.

```r
## Compute a 95% CI assuming that the variances are unknown and unequal
A <- c(25.6,23.2,26.5,25.8,24.7,23.8,25.3,24.1,24.2,26.2)
B <- c(29,29.3,31.1,29.1,29.5,29.6,27.6,27.9,28.5,30)

library(DescTools)
t.test(A, B, paired=FALSE, var.equal=FALSE, conf.level=0.95)
```

```
##
##  Welch Two Sample t-test
##
## data:  A and B
## t = -8.897, df = 17.877, p-value = 5.529e-08
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -5.216998 -3.223002
## sample estimates:
## mean of x mean of y
##     24.94     29.16
```

2a.

```r
library(DescTools)
muA <- 25
nA <- 12
varA <- 5

muB <- 40
nB <- 15
varB <- 5

mu_diff <- muA-muB
mu_diff
```

```
## [1] -15
```

```r
alpha <- 0.05
num_simulations <- 1000
coverage_count <- 0

for (i in 1:num_simulations) {
  sample_A <- rnorm(nA, mean = muA, sd = sqrt(varA))
  sample_B <- rnorm(nB, mean = muB, sd = sqrt(varB))

  results <- t.test(sample_A, sample_B, var.equal = TRUE, conf.level = 1 - alpha)

  if (mu_diff >= results$conf.int[1] && mu_diff <= results$conf.int[2]) {
    coverage_count <- coverage_count + 1
  }
}
coverage_rate <- coverage_count / num_simulations
coverage_rate
```

```
## [1] 0.953
```

The coverage rate for this simulation was 0.947. This is very close to the 1-α rate of 0.95, and is equal when rounded to the nearest hundreth.

2b.

```r
library(DescTools)
muA <- 25
nA <- 12
varA <- 9

muB <- 40
nB <- 15
varB <- 5

mu_diff <- muA-muB
mu_diff
```

```
## [1] -15
```

```r
alpha <- 0.05
num_simulations <- 1000
coverage_count <- 0

for (i in 1:num_simulations) {
  sample_A <- rnorm(nA, mean = muA, sd = sqrt(varA))
  sample_B <- rnorm(nB, mean = muB, sd = sqrt(varB))

  results <- t.test(sample_A, sample_B, var.equal = TRUE, conf.level = 1 - alpha)

  if (mu_diff >= results$conf.int[1] && mu_diff <= results$conf.int[2]) {
    coverage_count <- coverage_count + 1
  }
}
coverage_rate <- coverage_count / num_simulations
coverage_rate
```

```
## [1] 0.938
```

The coverage rate is 0.944. This is also surprisingly, very close to 0.95. This shows that making an incorrect assumption about if the variances are equal is not detrimental to the confidence interval.