Stat 3202 Lab 8

Nathan Johnson.9254

2023-03-31

Problem 1: Part a:

$$(\bar{x} - \bar{y}) \pm z_{1-\alpha/2}\sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}$$

Part b:

```
set.seed(1)

mux = 10
sx = 3
nx = 5
muy = 5
sy = 5.5
ny = 8
alpha = 0.05

coverage = c()

for(i in 1:30000) {
  samplex = rnorm(nx, mux, sx)
  sampley = rnorm(ny, muy, sy)
  x = mean(samplex)
  y = mean(sampley)

  z = qnorm(1-alpha/2)

  CIlower = (x - y) - z * sqrt(sx^2/nx + sy^2/ny)
  CIUpper = (x - y) + z * sqrt(sx^2/nx + sy^2/ny)

  coverage[i] = CIlower <= mux - muy && mux - muy <= CIUpper
}

mean(coverage)
```

```
## [1] 0.9481333
```

The coverage rate for this confidence interval is 95% which is what we expect.

Part c: Part i.

```r
set.seed(1)

mux = 10
sx = 3
nx = 5
muy = 5
sy = 3
ny = 8
alpha = 0.05

t = qt(1-alpha/2, nx+ny-2)

coverage = c()

for(i in 1:30000) {
  samplex = rnorm(nx, mux, sx)
  sampley = rnorm(ny, muy, sy)
  x = mean(samplex)
  y = mean(sampley)

  s2x = var(samplex)
  s2y = var(sampley)

  s2p = ((nx - 1)*s2x + (ny - 1)*s2y) / (nx + ny - 2)

  CIlower = (x - y) - t * sqrt(s2p) * sqrt(1/nx + 1/ny)
  CIUpper = (x - y) + t * sqrt(s2p) * sqrt(1/nx + 1/ny)

  coverage[i] = CIlower <= mux - muy && mux - muy <= CIUpper
}

mean(coverage)
```

```
## [1] 0.9500333
```

The coverage rate for this confidence interval is 95% which is what we expect.

Part ii.

```r
set.seed(1)

mux = 10
sx = 3
nx = 5
muy = 5
sy = 3
ny = 8
alpha = 0.05

coverage = c()

for(i in 1:30000) {
  samplex = rnorm(nx, mux, sx)
```

```
  sampley = rnorm(ny, muy, sy)

  output = t.test(samplex, sampley, conf.level = 1-alpha, var.equal = TRUE)

  CIlower = output$conf.int[1]
  CIupper = output$conf.int[2]

  coverage[i] = CIlower <= mux - muy && mux - muy <= CIUpper
}

mean(coverage)
```

```
## [1] 0.9748667
```

Despite copying Mr. Gan's work precisely, I can't seem to get this to run correctly. But the confidence interval should be 95%.

Problem 2: Part a: Part i.

$$v = \frac{(\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y})^2}{\frac{1}{n_x - 1}(\frac{s_x^2}{n_x}) + \frac{1}{n_y - 1}(\frac{s_y^2}{n_y})}$$

Part ii.

```
set.seed(1)

mux = 10
sx = 3
nx = 5
muy = 5
sy = 3
ny = 8
alpha = 0.05

coverage = c()

for(i in 1:30000) {
  samplex = rnorm(nx, mean = mux, sd = sx)
  sampley = rnorm(ny, mean = muy, sd = sy)

  output = t.test(samplex, sampley, conf.level = 1-alpha, var.equal = FALSE)

  CIlower = output$conf.int[1]
  CIupper = output$conf.int[2]

  coverage[i] = CIlower <= (mux - muy) && (mux - muy) <= CIUpper
}

mean(coverage)
```

```
## [1] 0.9764667
```

Once again, it didn't seem to compute the right values, but I believe my code to be correct. The confidence interval should have a confidence of 95%.

Problem 3:

```r
set.seed(1)

mux = 10
sx = 3
nx = 5
muy = 5
sy = 3
ny = 8
alpha = 0.05

v = min(nx - 1, ny - 1)
t = qt(1-alpha/2, v)

coverage = c()

for(i in 1:30000) {
  samplex = rnorm(nx, mean = mux, sd = sx)
  sampley = rnorm(ny, mean = muy, sd = sy)

  x = mean(samplex)
  y = mean(sampley)

  CIlower = (x - y) - t * sqrt((sx^2/nx)+(sy^2/ny))
  CIupper = (x - y) + t * sqrt((sx^2/nx)+(sy^2/ny))

  coverage[i] = CIlower <= (mux - muy) && (mux - muy) <= CIUpper
}

mean(coverage)
```

```
## [1] 0.9969667
```

The coverage rate is 99.5% which does show that the margin of error is higher and gives a greater coverage rate than 95%.