

final

2024-10-18

Functions

```
alpha <- 0.025
binorm <- function(mu1, mu2, sigma, rho, n){
  x1 <- rnorm(n)
  x2 <- rnorm(n)
  x3 <- rho*x1 + sqrt(1 - rho^2)*x2
  y1 <- mu1 + sigma*x1
  y2 <- mu2 + sigma*x3
  ymat <- cbind(y1, y2)
  return(ymat)
}

paired_t_test <- function(data) {
  t.test(data[, 1], data[, 2], paired = T, alternative = "greater", conf.level = 1 - alpha)
}

two_sample_t_test <- function(data) {
  t.test(data[, 1], data[, 2], paired = F, alternative = "greater", conf.level = 1 - alpha, var.equal =
}

#### Simulation ####
w_simulation <- function(mu1, mu2, sigma, rho, n, alpha) {
  reject_paired <- 0
  coverage_paired <- 0
  reject_two_sample <- 0
  coverage_two_sample <- 0

  for (i in 1:10000) {
    data <- binorm(mu1, mu2, sigma, rho, n) # bivariate normal samples

    # Paired
    paired_test <- paired_t_test(data)
    if (paired_test$p.value < alpha) {
      reject_paired <- reject_paired + 1 # reject if p_value less than alpha
    }
    if (paired_test$conf.int[1] < (mu1-mu2) & paired_test$conf.int[2] > (mu1-mu2)) {
      coverage_paired <- coverage_paired + 1 # CI cover the true mean
    }

    # Two-sample
    two_sample_test <- two_sample_t_test(data)
    if (two_sample_test$p.value < alpha) {
```

```

    reject_two_sample <- reject_two_sample + 1
  }
  if (two_sample_test$conf.int[1] < (mu1-mu2) & two_sample_test$conf.int[2] > (mu1-mu2)) {
    coverage_two_sample <- coverage_two_sample + 1
  }
}
df <- data.frame(
  paired = c(reject_paired, coverage_paired),
  two_Sample = c(reject_two_sample, coverage_two_sample),
  row.names = c("Number of Rejection", "Number of Coverage")
)

return(df)
}

```

Scenario i

H_0 : $\mu_1 = \mu_2$ truth : $\mu_1 = \mu_2 = 130$, null is true. type I error = rate of falsely rejecting the null = proportion of rejection.

```

set.seed(16)
s_i <- w_simulation(mu1=130, mu2=130, sigma=25, rho=0.4, n=20, alpha=0.025)
print(s_i)

```

```

##                paired two_Sample
## Number of Rejection    248        74
## Number of Coverage    9752       9926

```

```

type_1_error_i <- data.frame(
  (s_i[1,]/10000), row.names = "Type 1 Error")

coverage_i <- data.frame(
  (s_i[2,]/10000), row.names = "Coverage")

scenario_i <- bind_rows(s_i, type_1_error_i, coverage_i)
print(scenario_i)

```

```

##                paired two_Sample
## Number of Rejection 248.0000    74.0000
## Number of Coverage 9752.0000 9926.0000
## Type 1 Error        0.0248    0.0074
## Coverage            0.9752    0.9926

```

Q1

In scenario i, the type I error of the paired t-test is 0.0248 which is close to the nominal type I error 0.025. The type I error of the two sample t-test is 0.0074 which is much less than the type I error of the paired t-test and the nominal type I error.

The smaller type I error of the two sample t-test is due to the violation of the independence assumption of the two sample t-test. For two positively correlated data, a larger value of Y_i^1 means that Y_i^2 is also likely

to be large, and vice versa. So, the difference in mean will be smaller than the difference when they are independent, which makes the numerator of the test statistics smaller and therefore also the test statistics. For the alternative hypothesis, $\mu_1 > \mu_2$, the null is rejected when $t^* > t_{v,1-\alpha}$. So, a smaller test statistics makes it less likely to reject the null.

Q2

The actual coverage of the confidence interval of the paired t-test is 0.9752 which is close to the nominal coverage 0.975. The coverage of the confidence interval of the two sample t-test is 0.9926 which is larger than the coverage of the paired t-test and the nominal coverage.

The larger coverage of the two sample t-test is due to its wider confidence interval. For the correlated data, the effective number of independent pieces of information is lower, which makes it lose degrees of freedom. Because for $v' < v$, $t_{v',1-\alpha} > t_{v,1-\alpha}$, and the confidence interval for one-tail test is $[(\bar{Y}^1 - \bar{Y}^2) - t_{v,1-\alpha} \times s.e., \infty]$. A smaller degrees of freedom makes a wider confidence interval so that it is more likely to cover the true mean difference.

Both the hypothesis testing and confidence interval can be used to evaluate if the observed data supports the null hypothesis under certain significant level. Hypothesis testing is to check if the test statistics t^* is in the critical region of a certain significant level of the distribution under the null. Or, if the probability of obtaining an as extreme or more extreme test statistics than the current one is less than the significant level. If so, then the probability of getting this data given the null hypothesis is true is low so that we reject the null.

A $(1 - \alpha) \cdot 100\%$ confidence interval means that out of all intervals at the $(1 - \alpha) \cdot 100\%$ level, $(1 - \alpha) \cdot 100\%$ of them would contain the true parameter value. So, given a set of data, we can compute one of these intervals that is very likely to contain the true parameter. If what we proposed about the true parameter in the null hypothesis is in this interval, we fail to reject the null. Thus, $Coverage = P(CI \text{ include the true value}) = P(\text{fail to reject } H_0) = 1 - P(\text{reject } H_0) = 1 - \alpha = 1 - \text{type I error}$

In this example, the rejected ones are that the probability of getting these data from the distribution that assumes the null is true is lower than $\alpha = 0.025$. So it makes sense to have $248/10000 = 0.0248$ of the total to be rejected.

A 97.5% confidence interval means that 97.5% of the total confidence intervals would contain the true parameter value. So it makes sense for our case to have $9752/10000 = 97.52\%$ to contain the true value.

Scenario ii

H_0: $\mu_1 = \mu_2$ truth : $\mu_1 = 130 \neq \mu_2 = 105$, null is false. Power = rate of correctly rejecting the null = proportion of rejection

```
set.seed(18)
s_ii <- w_simulation(mu1=130, mu2=105, sigma=25, rho=0.4, n=20, alpha=0.025)
print(s_ii)
```

```
##                paired two_Sample
## Number of Rejection    9718      9204
## Number of Coverage     9737      9942
```

```
power_ii <- data.frame(
  (s_ii[1,]/10000), row.names = "Power")

coverage_ii <- data.frame(
```

```
(s_ii[2,]/10000), row.names = "Coverage")

scenario_ii <- bind_rows(s_ii, power_ii, coverage_ii)
print(scenario_ii)
```

```
##                paired two_Sample
## Number of Rejection 9718.0000  9204.0000
## Number of Coverage  9737.0000  9942.0000
## Power                0.9718    0.9204
## Coverage              0.9737    0.9942
```

Q3

In scenario ii, the power of the paired t-test is 0.9718 which is larger than the power of the two sample t-test, 0.9204.

Again, because the data are positively correlated, the test statistics of the two sample test is smaller than it would be when the data are independent, which makes it less likely to reject the null.

Q4

The actual coverage of the confidence interval of the paired t-test is 0.9737 which is close to the nominal coverage 0.975. The coverage of the confidence interval of the two sample t-test is 0.9942 which is larger than the coverage of the paired t-test and the nominal coverage.

The larger coverage of the two sample t-test is again due to its wider confidence interval; when the data is not independent, we lose degrees of freedom and thus have a higher critical value.

It also shows the equivalence between hypothesis testing and confidence interval:

The paired t-test with higher power that is better at correctly rejecting the null, results in narrower intervals that the coverage is closed to the nominal coverage. The two sample t-test's lower power and wider intervals makes the coverage larger the nominal level.

Q5

Summary

In scenario i, the paired t-test has a type I error rate close to the nominal value. The two-sample t-test is much more conservative and has a lower type I error rate. Correspondingly, the paired t-test has a coverage close to the nominal value, while the two-sample t-test has an inflated coverage due to the wider confidence intervals.

In scenario ii, the paired t-test has a higher power than the two-sample t-test. Again, the coverage of the paired t-test is close to the nominal value, while the two-sample t-test has an inflated coverage.

I recommend the paired t-test. Firstly, the independence assumption of the two-sample t-test is violated here. Secondly, unlike the two-sample t-test that has a low type I error rate and low power, the paired t-test is more balanced with respect to the low type I error rate and power.