

Power Study to Compare Mean Vectors from Two Populations

In order to compare mean vectors of two independent populations, we make the following assumptions:

1. $X_{1,1}, \dots, X_{1,n_1}$ is a random sample of size n_1 from a p -variate population with mean vector μ_1 and covariance matrix Σ_1 .
2. $X_{2,1}, \dots, X_{2,n_2}$ is a random sample of size n_2 from a p -variate population with mean vector μ_2 and covariance matrix Σ_2 .
3. $\{X_{1,1}, \dots, X_{1,n_1}\}$ is independent of the other $\{X_{2,1}, \dots, X_{2,n_2}\}$.

We want to test for the equality of the mean vectors from two populations:

$$H_0: \mu_1 = \mu_2 \text{ vs. } H_a: \mu_1 \neq \mu_2$$

If we add two more assumptions:

4. $X_{1,1}, \dots, X_{1,n_1} \sim \text{MN}(\mu_1, \Sigma_1)$ and $X_{2,1}, \dots, X_{2,n_2} \sim \text{MN}(\mu_2, \Sigma_2)$
5. $\Sigma_1 = \Sigma_2$

Then, the likelihood ratio test of H_0 is based on the square of the statistical distance, T^2 , and is given by:

$$T^2 = (X_1 - X_2 - (\mu_1 - \mu_2))' \left(\frac{1}{n_1} + \frac{1}{n_2} \right) S_{pooled}^{-1} (X_1 - X_2 - (\mu_1 - \mu_2))$$

where:

- T^2 is distributed as $\frac{(n_1+n_2-2)p}{n_1+n_2-p-1} F_{p, n_1+n_2-p-1}(\alpha)$
- $S_{pooled} = \frac{\sum_{j=1}^{n_1} (X_{1j} - X_1)(X_{1j} - X_1)' + \sum_{j=1}^{n_2} (X_{2j} - X_2)(X_{2j} - X_2)'}{n_1+n_2-2} = \frac{n_1-1}{n_1+n_2-2} S_1 + \frac{n_2-1}{n_1+n_2-2} S_2$
- the critical distance c^2 is determined from the distribution of the two-sample T^2 statistic so that $P(T^2 < c^2) = \frac{(n_1+n_2-2)p}{n_1+n_2-p-1} F_{p, n_1+n_2-p-1}(\alpha) = 1 - \alpha$.

Therefore, we reject H_0 if: $T^2 > c^2$. We conclude that all $\mu_1 - \mu_2$ within squared statistical distance c^2 of $x_1 - x_2$ constitute the confidence region.

When the assumptions (4) and/or (5) is/are not met, for large samples, this structure is sufficient for making inferences about the $p \times 1$ vector $H_0: \mu_1 = \mu_2$. The approximate test that allow us to test the hypotheses $H_0: \mu_1 = \mu_2$ vs $H_a: \mu_1 \neq \mu_2$ is:

$$T^2 = (X_1 - X_2 - (\mu_1 - \mu_2))' \left(\frac{1}{n_1} S_1 + \frac{1}{n_2} S_2 \right) (X_1 - X_2 - (\mu_1 - \mu_2)) \sim \chi_p^2.$$

We reject H_0 at level α if $T^2 > \chi_{p,\alpha}^2$.

We construct a power study that estimates the power and type-I error of the approximate test when assumptions 4 and 5 are not met.

1) Testing when Assumption (4) is not met

We test $H_0: \mu_1 = \mu_2 = \mu_0 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ when:

- 4. $X_{1,1}, \dots, X_{1,n_1} \sim \text{MN}(\mu_1, \Sigma_1)$ and $X_{2,1}, \dots, X_{2,n_2} \sim \text{MN}(\mu_2, \Sigma_2)$ is not met.

In order to equality of the mean, we vary the third component of μ_0 between -2 and 2 i.e we test for

$$\mu_0 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \mu_3 \end{pmatrix} \text{ with } \mu_3 \text{ varying between -2 and 2 at an interval of 0.2.}$$

We compare the Mean Vectors from Two Populations with different distributions:

1. normal distribution vs normal distribution
2. skew-normal distribution vs skew-normal distribution
3. skew-t distribution vs skew-t distribution
4. normal distribution vs skew-normal distribution
5. normal distribution vs skew-t distribution
6. skew-normal distribution vs skew-t distribution

Remarks:

When n grows, the power performs better on the three tests (normal vs normal, normal vs skew-normal and normal vs skew-t). We can see that around 0, the power shrinks as the sample size n is bigger.

```
## # A tibble: 18 x 4
## # Groups:   mu, test [6]
##      mu test      n    power
##    <dbl> <chr>    <fct> <dbl>
##  1      0 reject_mv_mv 10     0.05
##  2      0 reject_mv_sn 10     0.06
##  3      0 reject_mv_st 10     0.065
##  4      0 reject_sn_sn 10     0.055
##  5      0 reject_sn_st 10     0.035
##  6      0 reject_st_st 10     0.05
##  7      0 reject_mv_mv 100    0.015
##  8      0 reject_mv_sn 100    0.01
##  9      0 reject_mv_st 100    0.01
## 10      0 reject_sn_sn 100    0.015
## 11      0 reject_sn_st 100    0.01
## 12      0 reject_st_st 100    0.005
## 13      0 reject_mv_mv 1000   0.005
## 14      0 reject_mv_sn 1000   0.015
## 15      0 reject_mv_st 1000   0.01
## 16      0 reject_sn_sn 1000   0.01
## 17      0 reject_sn_st 1000   0.01
## 18      0 reject_st_st 1000   0.01
```

However, for $n = 10$, the estimated type-I error is very high ranging between with .035 and 0.065 for a significance level $\alpha = 0.01$. This result makes sense because the test follows approximately a χ^2 -distribution only when the two samples sizes are large. We can also see that the mean of the power for $n = 100$ is equal to 0.0108 which is higher than the significance level α . Therefore, a large sample in this case should be at least bigger than 10 in order to be accurate.

```
## # A tibble: 3 x 2
##      n    mean_power
##    <fct>    <dbl>
## 1 10      0.0525
## 2 100     0.0108
## 3 1000    0.01
```

When comparing the power of the different distributions, we can see that the power is higher when comparing two multivariate normal distributions. We can also see that the test performs better on the skew-normal compared to the skew-t because the skew-normal is more similar to the normal distribution (compared to the skew-t distribution). When comparing the type-I error of the different distributions, we can see that the 6 different tests are similar with an error approximately equal to 0.02 (or a little bit higher than the significance level $\alpha = 0.01$).

```
## # A tibble: 6 x 2
##      test    mean_power
##    <chr>    <dbl>
## 1 reject_mv_mv 0.0233
## 2 reject_mv_sn 0.0283
## 3 reject_mv_st 0.0283
## 4 reject_sn_sn 0.0267
## 5 reject_sn_st 0.0183
## 6 reject_st_st 0.0217
```

2) Testing when Assumption (5) is not met

We test $H_0: \mu_1 = \mu_2 = \mu_0 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ when:

- 5. $\Sigma_1 = \Sigma_2$ is not met.

In order to equality of the mean, we vary the third component of μ_0 between -2 and 2 i.e we test for

$$\mu_0 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \mu_3 \end{pmatrix} \text{ with } \mu_3 \text{ varying between -2 and 2 at an interval of 0.2.}$$

We compare the Mean Vectors from two normal populations with different covariance matrix.

In order to test for the covariance matrices of the two normally distributed populations, we test for

$$\Sigma_1 = \Sigma_2 = \Sigma_0 = \begin{pmatrix} 1 & 0.5 & 0.5 \\ 0.5 & 1 & 0.5 \\ 0.5 & 0.5 & 1 \end{pmatrix}.$$

We control the distribution of the second covariance matrix by multiplying it with an integers between 1 and 10. Therefore, we can choose k such that $\Sigma_2 = k\Sigma_1 = k\Sigma_0$. For k = 1, the two covariance matrices are equal. When k gets bigger, we have larger violation of assumption 5.

Remarks:

When n grows, the power performs better on the six tests. We can see that around 0, the power shrinks around as the sample size n is bigger.

```
## # A tibble: 30 x 4
## # Groups:   mu, sigma [10]
##       mu sigma n      power
##   <dbl> <dbl> <fct> <dbl>
## 1     0     1 10    0.055
## 2     0     2 10    0.045
## 3     0     3 10    0.06
## 4     0     4 10    0.06
## 5     0     5 10    0.055
## 6     0     6 10    0.07
## 7     0     7 10    0.075
## 8     0     8 10    0.1
## 9     0     9 10    0.07
## 10    0    10 10    0.065
## # ... with 20 more rows
```

However, for $n = 10$, the estimated type-I error is very high ranging between with 0.045 and 0.1 for a significance level $\alpha = 0.01$. This result makes sense because the test follows approximately a χ^2 -distribution only when the two samples sizes are large. We can also see that the mean of the power for $n = 100$ or $n = 1000$ is maintained approximately around 0.01. Therefore, a large sample in this case should be at least bigger than 10 in order to be accurate.

```
## # A tibble: 3 x 2
##   n      mean_power
##   <fct>         <dbl>
## 1 10          0.0655
## 2 100         0.0155
## 3 1000        0.0115
```

When comparing the power of the test for the two population means, we can see that as k gets bigger, the power around the mean 0 grows. I.e., when the integer multiplying Σ_0 is closer to 1, the power around the mean 0 shrinks. When the integer multiplying Σ_0 is closer to 10, the power around the mean 0 grows. We can conclude

that for smaller violation of assumption 5, the power is higher when $\mu_2 \neq \mu_0 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$. For larger violation of

assumption 5, the power is lower when $\mu_2 \neq \mu_0 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$. When comparing the type-I error of the test for the two

population means, we can see that the error is approximately equal to 0.03 (or higher than the significance level $\alpha = 0.01$).

```
## # A tibble: 10 x 2
##   sigma mean_power
##   <dbl>     <dbl>
## 1     1     0.0283
## 2     2     0.0233
## 3     3     0.03
## 4     4     0.0233
## 5     5     0.0283
## 6     6     0.0333
## 7     7     0.0350
## 8     8     0.05
## 9     9     0.0283
## 10    10     0.0283
```

3) Testing when Assumption (4) and Assumption (5) are not met

We test $H_0: \mu_1 = \mu_2 = \mu_0 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ when:

- 4. $X_{1,1}, \dots, X_{1,n_1} \sim \text{MN}(\mu_1, \Sigma_1)$ and $X_{2,1}, \dots, X_{2,n_2} \sim \text{MN}(\mu_2, \Sigma_2)$ is not met.
- 5. $\Sigma_1 = \Sigma_2$ is not met.

In order to equality of the mean, we vary the third component of μ_0 between -2 and 2 i.e we test for

$$\mu_0 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \mu_3 \end{pmatrix} \text{ with } \mu_3 \text{ varying between -2 and 2 at an interval of 0.2.}$$

We compare the Mean Vectors from two topulations with different distributions and with different covariance matrices.

The couples of distributions are:

1. normal distribution vs normal distribution
2. normal distribution vs skew-normal distribution
3. normal distribution vs skew-t distribution

(We do not test for all the combinations as the function would take too long to run).

We control the distribution of the second covariance matrix by multiplying it with an integers between 1 and 10. Therefore, we can choose k such that $\Sigma_2 = k\Sigma_1 = k\Sigma_0$. For $k = 1$, the two covariance matrices are equal. When k gets bigger, we have larger violation of assumption 5.

Remarks:

We observe the same remarks that given previously:

When n grows, the power performs better on the 4 different tests of the mean comparison for any covariance matrix. We can see that around 0, the power shrinks as the sample size n is bigger.

```
## # A tibble: 120 x 5
## # Groups:   mu, sigma, test [40]
##      mu sigma test      n    power
##    <dbl> <dbl> <chr>    <fct> <dbl>
##  1      0      1 reject_mv_mv 10    0.06
##  2      0      1 reject_mv_sn 10    0.04
##  3      0      1 reject_mv_st 10    0.045
##  4      0      1 reject_sn_st 10    0.035
##  5      0      2 reject_mv_mv 10    0.04
##  6      0      2 reject_mv_sn 10    0.06
##  7      0      2 reject_mv_st 10    0.08
##  8      0      2 reject_sn_st 10    0.055
##  9      0      3 reject_mv_mv 10    0.06
## 10      0      3 reject_mv_sn 10    0.09
## # ... with 110 more rows
```

However, for $n = 10$, the estimated type-I error is very high with an average of 0.060375 for a significance level $\alpha = 0.01$. This result makes sense because the test follows approximately a χ^2 -distribution only when the two samples sizes are large. We can also see that the mean of the power for $n = 100$ or $n = 1000$ is maintained approximately around 0.01. Therefore, a large sample in this case should be at least bigger than 10 in order to be accurate.

```
## # A tibble: 3 x 2
##   n    mean_power
##   <fct>      <dbl>
## 1 10      0.0604
## 2 100     0.0115
## 3 1000    0.0106
```

When comparing the power of the different distributions, we can see that the power is higher when comparing two multivariate normal distributions. We can also see that the test performs better on the skew-normal compared to the skew-t because the skew-normal is more similar to the normal distribution (compared to the skew-t distribution). When comparing the type-I error of the different distributions, we can see that the 4 different tests are similar with an error approximately equal to 0.028 (or higher than the significance level $\alpha = 0.01$).

```
## # A tibble: 4 x 2
##   test      mean_power
##   <chr>      <dbl>
## 1 reject_mv_mv    0.0285
## 2 reject_mv_sn    0.0293
## 3 reject_mv_st    0.0283
## 4 reject_sn_st    0.0238
```

When comparing the power of the test for the two population means, we can see that as k gets bigger, the power around the mean 0 grows. I.e., when the integer multiplying Σ_0 is closer to 1, the power around the mean 0 shrinks. When the integer multiplying Σ_0 is closer to 10, the power around the mean 0 grows. We can conclude

that for smaller violation of assumption 5, the power is higher when $\mu_2 \neq \mu_0 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$. For larger violation of

assumption 5, the power is lower when $\mu_2 \neq \mu_0 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$. When comparing the type-I error of the test for the two

population means, we can see that the error is approximately equal to 0.03 (or higher than the significance level $\alpha = 0.01$).

```
## # A tibble: 10 x 2
##   sigma mean_power
##   <dbl>      <dbl>
## 1     1    0.02
## 2     2   0.0279
## 3     3   0.0275
## 4     4   0.0258
## 5     5    0.02
## 6     6    0.03
## 7     7   0.0333
## 8     8   0.0308
## 9     9   0.0304
## 10    10   0.0292
```