MATH 189
Global Mean Testing and Two-Sample Problems

Wenxin Zhou
UC San Diego

Time: 2:00—3:20 & 3:30—4:50pm TueThur
Location: CENTR 115

Outline

- In the last lecture, we considered multiple testing problems for population mean vector.
  - Multiple testing
  - Type I error control: FWER control, FDR control
  - Methods: Bonferroni correction, Holm's method, BH-procedure

- Today we will discuss more hypothesis testing problems on population mean vectors.
  - Hotelling's T^2
  - Paired samples
  - Two-sample problems

Go Beyond Multiple Testing

- Thus far the mean vector testing schemes we have discussed are based on comparing each component of two vectors.
- The problem of testing the equivalence of two vectors is decomposed to testing multiple univariate hypotheses.
- This type of testing scheme is called multiple testing.
- There are other ways to test the mean vector from one or multiple populations.

Hotelling's T^2 (Univariate)

- Consider the square of the t-statistic for testing a univariate mean:

$$T^2 = \frac{(\bar{x} - \mu)^2}{s^2/n} = n(\bar{x} - \mu)\frac{1}{s^2}(\bar{x} - \mu) \sim F_{1,n-1}$$

- Squaring a t-distributed random variable with $n - 1$ degrees of freedom gives rise to an F-distributed random variable with 1 and $n - 1$ degrees of freedom.

- We reject H_0 at level α if T^2 exceeds the critical value from the F-table with 1 and $n - 1$ degrees of freedom, evaluated at level α:

$$T^2 > F_{1,n-1,\alpha}$$

Hotelling's T^2 (Multivariate)

- Motivated by the univariate case, we define the Hotelling's T^2 statistic for testing a mean vector

$$T^2 = n(\overline{\boldsymbol{x}} - \boldsymbol{\mu}_0)'\mathbf{S}^{-1}(\overline{\boldsymbol{x}} - \boldsymbol{\mu}_0)$$

- The difference between the sample mean and μ_0 is replaced with the difference between the sample mean vector and the hypothesized mean vector $\boldsymbol{\mu}_0$.
- The inverse of the sample variance is replaced by the inverse of the sample covariance matrix $\mathbf{S}$.

Distribution of Hotelling's T^2 (Large Sample)

- When sample size n is large, we can ignore the estimation error of variance-covariance matrix.
- If we replace the sample covariance matrix $\mathbf{S}$ by the population covariance matrix $\mathbf{\Sigma}$, then T^2 is chi-square distributed with m degrees of freedom.

$$n(\bar{\boldsymbol{x}} - \boldsymbol{\mu}_0)'\mathbf{\Sigma}^{-1}(\bar{\boldsymbol{x}} - \boldsymbol{\mu}_0) \sim \chi^2_{\mathrm{m}}$$

- Therefore, when the sample size n is large, T^2 is *approximately* chi-square distributed with m degrees of freedom.

Distribution of Hotelling's T^2 (Small Sample)

- For small samples, the chi-square approximation for T^2 does not take into account the estimation error of estimating $\mathbf{\Sigma}$ with $\mathbf{S}$.
- Better results can be obtained from the following transformation of the Hotelling's T^2 statistic:

$$F = \frac{n-m}{m(n-1)} T^2 \sim F_{m,n-m}$$

- Under H_0, F follows a F-distribution with m and $n-m$ degrees of freedom.
- We reject the null hypothesis at level α if

$$F > F_{m,n-m,\alpha}$$

Example: USDA Women's Health Survey Data

- The recommended intake and sample mean are given below

Variable	Recommended Intake (μ_0)	Sample Mean ($\bar{x}$)
Calcium	1000 mg	624.0 mg
Iron	15mg	11.1 mg
Protein	60g	65.8 g
Vitamin A	800 µg	839.6 µg
Vitamin C	75 mg	78.9 mg

At a significance level $\alpha = 0.01$, test the following null and alternative hypotheses using Hotelling's T^2:

$$H_0: \boldsymbol{\mu} = \boldsymbol{\mu}_0 \quad \text{vs} \quad H_1: \boldsymbol{\mu} \neq \boldsymbol{\mu}_0$$

Example: USDA Women's Health Survey Data

- Calculate sample covariance matrix $\boldsymbol{S}$:

	Calcium	Iron	Protein	Vitamin A	Vitamin C
Calcium	157829.4	940.1	6075.8	102411.1	6701.6
Iron	940.1	35.8	114.1	2383.2	137.7
Protein	6075.8	114.1	934.9	7330.1	477.2
Vitamin A	102411.1	2383.2	7330.1	2668452.4	22063.3
Vitamin C	6701.6	137.7	477.2	22063.3	5416.3

Example: USDA Women's Health Survey Data

- The Hotelling's T^2 can be calculated as

$$T^2 = n(\bar{\boldsymbol{x}} - \boldsymbol{\mu}_0)'\mathbf{S}^{-1}(\bar{\boldsymbol{x}} - \boldsymbol{\mu}_0) = 1758.54.$$

- The F-statistic can be calculated as

$$F = \frac{n-m}{m(n-1)}T^2 = 349.80.$$

The critical value can be obtained from F-table (or by R) with m and $n - m$ degrees of freedom, evaluated at level α ($n = 737, m = 5, \alpha = 0.01$):

$$F_{5,732,0.01} = 3.042.$$

Conclusion: We reject the null as the F-statistic exceeds the critical value.

Compare Two Population Means

- Thus far, we have focused on testing if a population mean vector equals a specific vector:

$$H_0: \boldsymbol{\mu} = \boldsymbol{\mu}_0 \quad \text{vs} \quad H_1: \boldsymbol{\mu} \neq \boldsymbol{\mu}_0.$$

- In some applications, we want to test the equivalence of two (unknown) population means, say

$$H_0: \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 \quad \text{vs} \quad H_1: \boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2,$$

  where $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ are the mean vectors of two populations.

- This is a two-sample testing problem, which will be our focus today.

Two-Sample Mean Testing Problems

- Two-sample mean testing has many applications:
    1. Test if there is a significant difference before and after a treatment.
    2. Compare the reading ability of the students in two middle schools.
    3. Check if a policy has a positive impact on reducing unemployment rate.
    4. Can we classify two cities into the same class according to their weather conditions?
    5. Which mobile phone company has better quality control? Samsung or Apple?
    6. Etc.

- There are several aspects we need to consider before the test:
    1. The data may either be paired or unpaired.
    2. The sample sizes can be the same or different.
    3. The covariance matrices of two populations can be equal or unequal (we often do not know this in practice).

Paired Samples

- Paired samples are samples where natural or matched couplings occur. This generates a dataset in which each data point in one sample is uniquely paired to a data point in the second sample.

Sample 1

- n observations
- m variables
- $\boldsymbol{x}_1^{(1)}, \boldsymbol{x}_2^{(1)}, \dots, \boldsymbol{x}_n^{(1)}$
- Population parameters: $\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1$
- Sample statistics: $\overline{\boldsymbol{x}}_1, \mathbf{S}_1$

Sample 2

- n observations
- m variables
- $\boldsymbol{x}_1^{(2)}, \boldsymbol{x}_2^{(2)}, \dots, \boldsymbol{x}_n^{(2)}$
- Population parameters: $\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2$
- Sample statistics: $\overline{\boldsymbol{x}}_2, \mathbf{S}_2$

$$
\begin{array}{cccc}
\boldsymbol{x}_1^{(1)} & \boldsymbol{x}_2^{(1)} & \dots & \boldsymbol{x}_n^{(1)} \\
\downarrow & \downarrow & & \downarrow \\
\boldsymbol{x}_1^{(2)} & \boldsymbol{x}_2^{(2)} & \dots & \boldsymbol{x}_n^{(2)}
\end{array}
$$

Two samples are one-to-one paired

Paired Samples (cont.)

1. Pre-test/post-test samples in which a factor is measured before and after an intervention.
2. Cross-over trials in which individuals are randomized to two treatments and then the same individuals are crossed-over to the alternative treatment.
3. Matched samples, in which individuals are matched on personal characteristics such as age and sex.
4. Duplicate measurements on the same biological samples.

5. And any circumstance in which each data point in one sample is uniquely matched to a data point in the second sample.

Example: Spouse Data

- A sample of 30 husband-wife pairs are asked to respond to each of the following questions:
    1. What is the level of passionate love you feel for your partner?
    2. What is the level of passionate love your partner feels for you?
    3. What is the level of companionate love you feel for your partner?
    4. What is the level of companionate love your partner feels for you?

- Responses were recorded on the five-point scale:

  1 None at all, 2. Very little, 3. Some, 4. A great deal 5. Tremendous amount

Example: Spouse Data

Population/Sample 1: Husbands

- 30 observations
- 4 variables (points to 4 questions)
- $\boldsymbol{x}_1^{(1)}, \boldsymbol{x}_2^{(1)}, \dots, \boldsymbol{x}_{30}^{(1)}$
- $\boldsymbol{x}_i^{(1)} = \left(x_{i1}^{(1)}, x_{i2}^{(1)}, x_{i3}^{(1)}, x_{i4}^{(1)}\right)'$
- Population parameters: $\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1$
- Sample statistics: $\overline{\boldsymbol{x}}_1, \boldsymbol{S}_1$

Population/Sample 2: Wives

- 30 observations
- 4 variables (points to 4 questions)
- $\boldsymbol{x}_1^{(2)}, \boldsymbol{x}_2^{(2)}, \dots, \boldsymbol{x}_n^{(2)}$
- $\boldsymbol{x}_i^{(2)} = \left(x_{i1}^{(2)}, x_{i2}^{(2)}, x_{i3}^{(2)}, x_{i4}^{(2)}\right)'$
- Population parameters: $\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2$
- Sample statistics: $\overline{\boldsymbol{x}}_2, \boldsymbol{S}_2$

A husband in Sample 1 is uniquely paired to a wife in Sample 2

Two questions of interest:

1. Do the husbands respond to the questions in the same way as their wives?
2. Do the husbands and wives accurately perceive the responses of their spouses?

Question 1: Do the husbands respond to the questions in the same way as their wives?

- To answer this question is equivalent to answer if the population mean vector of the husbands population equals the population mean vector of the wives population.

$$\text{Yes: } \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 \text{ or } \text{No: } \boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2.$$

To answer this question, we are interested in testing the following null and alternative hypotheses:

$$H_0: \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 \quad \text{vs} \quad H_1: \boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2.$$

Testing on Paired Samples

- We can re-write the null and alternative hypotheses as

$$H_0: \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2 = \mathbf{0} \quad \text{vs} \quad H_1: \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2 \neq \mathbf{0}.$$

- This motivates us to consider a pair-wise difference of the two samples.
- Denote $\boldsymbol{y}_i$ the difference in the i-th couple (i-th husband and i-th wife)

$$\boldsymbol{y}_i = \boldsymbol{x}_i^{(1)} - \boldsymbol{x}_i^{(2)}.$$

- The population mean vector of $\boldsymbol{y}_i$ is

$$\boldsymbol{\mu}_y = \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2.$$

- So we transform the two-sample testing problem to a one-sample problem:

$$H_0: \boldsymbol{\mu}_y = \mathbf{0} \quad \text{vs} \quad H_1: \boldsymbol{\mu}_y \neq \mathbf{0}.$$

Paired Hotelling's T^2

- As $\boldsymbol{y}_i = \boldsymbol{x}_i^{(1)} - \boldsymbol{x}_i^{(2)}$ can be considered a new sample, we can calculate its sample mean vector and sample covariance matrix as

$$\overline{\boldsymbol{y}} = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{y}_i \quad \text{and} \quad \mathbf{S}_{\boldsymbol{y}} = \frac{1}{n-1}\sum_{i=1}^{n}(\boldsymbol{y}_i - \overline{\boldsymbol{y}})(\boldsymbol{y}_i - \overline{\boldsymbol{y}})'.$$

- The paired Hotelling's T^2 test statistic is given by

$$T^2 = n\overline{\boldsymbol{y}}'\mathbf{S}_{\boldsymbol{y}}^{-1}\overline{\boldsymbol{y}}.$$

- Then, we can define the F-statistic as

$$F = \frac{n-m}{m(n-1)}T^2 \sim F_{m,n-m}$$

At a given significance level α, we reject the null hypothesis $H_0\colon \boldsymbol{\mu}_y = \mathbf{0}$ if

$$F > F_{m,n-m,\alpha}$$

Question 1: Do the husbands respond to the questions in the same way as their wives?

1. Calculate sample mean vector

Question	Sample mean $\bar{y}$
1	0.0667
2	-0.1333
3	-0.3000
4	-0.1333

2. Calculate sample covariance matrix

$$\mathbf{S}_y = \begin{pmatrix} 0.8230 & 0.0782 & -0.0138 & -0.0598 \\ 0.0782 & 0.8092 & -0.2138 & -0.1563 \\ -0.0138 & -0.2138 & 0.5621 & 0.5103 \\ -0.0598 & -0.1563 & 0.5103 & 0.6023 \end{pmatrix}$$

3. T^2-statistic

$$T^2 = n\bar{\mathbf{y}}'\mathbf{S}_y^{-1}\bar{\mathbf{y}} = 13.13.$$

4. F-statistic

$$F = \frac{n-m}{m(n-1)}T^2 = 2.94.$$

5. Critical value ($\alpha = 0.05$)

$$F_{m,n-m,\alpha} = 2.74.$$

6. We reject the null hypothesis H_0: $\boldsymbol{\mu}_y = \mathbf{0}$ if

$$F > F_{m,n-m,\alpha}.$$

Question 2: Do the husbands and wives accurately perceive the responses of their spouses?

- To understand this question, let us return to the four questions asked of each husband-wife pair. The questions were:
  1. What is the level of passionate love you feel for your partner?
  2. What is the level of passionate love your partner feels for you?
  3. What is the level of companionate love you feel for your partner?
  4. What is the level of companionate love your partner feels for you?
- A sub-question of question 2 is:
  - Does the wife accurately perceive the amount of passionate love her husband feels towards her?
- This is equivalent to ask:
  - Does the husband's answer to question 1 match the wife's answer to question 2?
- Mathematically, we formulate the question into the following null and alternative hypotheses

$$H_0: \mu_{1,1} - \mu_{2,2} = 0 \quad \text{vs} \quad H_1: \mu_{1,1} - \mu_{2,2} \neq 0.$$

Question 2: Do the husbands and wives accurately perceive the responses of their spouses?

- To address question 2 we need to define four new variables as follows:
  1. $z_{i,1} = x_{i,1}^{(1)} - x_{i,1}^{(2)}$ (husbands' response to question 1 minus wives' response to question 2).
  2. $z_{i,2} = x_{i,2}^{(1)} - x_{i,1}^{(2)}$ (husbands' response to question 2 minus wives' response to question 1).
  3. $z_{i,3} = x_{i,3}^{(1)} - x_{i,4}^{(2)}$ (husbands' response to question 3 minus wives' response to question 4).
  4. $z_{i,4} = x_{i,4}^{(1)} - x_{i,3}^{(2)}$ (husbands' response to question 4 minus wives' response to question 3).

- Then, we transform the question 2 to the following mean testing problem:

$$H_0: \boldsymbol{\mu}_z = \mathbf{0} \quad \text{vs} \quad H_1: \boldsymbol{\mu}_z \neq \mathbf{0}.$$

- Use the paired Hotelling's T^2 test.

Conclusion for Question 2

- Calculate T^2-statistic

$$T^2 = n\bar{\mathbf{z}}'\mathbf{S}_{\mathbf{z}}^{-1}\bar{\mathbf{z}} = 6.43.$$

- Calculate F-statistic

$$F = \frac{n-m}{m(n-1)}T^2 = 1.44.$$

- Calculate critical value ($\alpha = 0.05$)

$$F_{m,n-m,\alpha} = 2.74.$$

- Decision: We do not reject the null hypothesis $H_0\colon \boldsymbol{\mu}_y = \mathbf{0}$ as

$$F \leq F_{m,n-m,\alpha}.$$

- Conclusion: There is no clear evidence, according to the observed sample, to reject the hypothesis: husbands and wives accurately perceive the responses of their spouses.

General Two-Sample Testing Problem

- In general, observations from the two samples may not be one-to-one paired. We may even observe two samples with very different sizes.

Sample 1

- n_1 observations
- m variables
- $\boldsymbol{x}_1^{(1)}, \boldsymbol{x}_2^{(1)}, \dots, \boldsymbol{x}_{n_1}^{(1)}$
- Population parameters: $\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1$
- Sample statistics: $\overline{\boldsymbol{x}}_1, \boldsymbol{S}_1$

Sample 2

- n_2 observations
- m variables
- $\boldsymbol{x}_1^{(2)}, \boldsymbol{x}_2^{(2)}, \dots, \boldsymbol{x}_{n_2}^{(2)}$
- Population parameters: $\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2$
- Sample statistics: $\overline{\boldsymbol{x}}_2, \boldsymbol{S}_2$

We are interested in testing the equivalence of the two population means:

$$H_0: \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 \quad \text{vs} \quad H_1: \boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2.$$

Two-Sample Hotelling's T^2

- The two-sample Hotelling's T^2 statistic is defined as

$$T^2 = (\bar{\boldsymbol{x}}_1 - \bar{\boldsymbol{x}}_2)' \left\{ \frac{1}{n_1} \mathbf{S}_1 + \frac{1}{n_2} \mathbf{S}_2 \right\}^{-1} (\bar{\boldsymbol{x}}_1 - \bar{\boldsymbol{x}}_2).$$

- Again, this is a function of the difference between the sample means for the two populations.
- Instead of being a function of a single covariance matrix, we can see that the two-sample test statistic is a function of $\mathbf{S}_1$, $\mathbf{S}_2$, n_1 and n_2.
- Denote $\mathbf{S}_{\mathrm{T}} = \left\{ \frac{1}{n_1} \mathbf{S}_1 + \frac{1}{n_2} \mathbf{S}_2 \right\}$.
- $\mathbf{S}_{\mathrm{T}}$ can be understood as a weighted-average of sample covariance matrices $\mathbf{S}_1$ and $\mathbf{S}_2$.

Two-Sample Hotelling's T^2 (Large Sample)

- When both n_1 and n_2 are large, we can ignore the effects of estimating the population covariance matrices by their sample estimators.

$$\mathbf{S}_1 \approx \boldsymbol{\Sigma}_1 \text{ and } \mathbf{S}_2 \approx \boldsymbol{\Sigma}_2.$$

- Therefore T^2 statistic is approximately chi-square distributed with degree of freedom m under null hypothesis.

$$T^2 \approx (\overline{\boldsymbol{x}}_1 - \overline{\boldsymbol{x}}_2)' \left\{\frac{1}{n_1}\boldsymbol{\Sigma}_1 + \frac{1}{n_2}\boldsymbol{\Sigma}_2\right\}^{-1} (\overline{\boldsymbol{x}}_1 - \overline{\boldsymbol{x}}_2) \sim \chi^2_{\mathrm{m}}.$$

- At a given significance level α, we reject the null hypothesis $H_0\colon \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$ if

$$T^2 > \chi^2_{\mathrm{m},\alpha}.$$

Two-Sample Hotelling's T^2 (Small Sample)

- When the sample sizes n_1 and n_2 are not large, we cannot ignore the effects of estimating the population covariance matrices by their sample counterparts.
- Instead, we can calculate an F transformation using the formula below

$$F = \frac{n_1+n_2-m-1}{m(n_1+n_2-2)} T^2 \sim F_{m,\nu}.$$

- The degree of freedom ν is defined through the following complicated formula

$$\frac{1}{\nu} = \sum_{i=1}^{2} \frac{1}{n_i-1} \left\{ \frac{(\bar{\boldsymbol{x}}_1-\bar{\boldsymbol{x}}_2)' \mathbf{S}_T^{-1} \left(\frac{1}{n_i} \mathbf{S}_i \right) \mathbf{S}_T^{-1} (\bar{\boldsymbol{x}}_1-\bar{\boldsymbol{x}}_2)}{T^2} \right\}^2.$$

- At a given significance level α, we reject the null hypothesis $H_0: \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$ if

$$F > F_{m,\nu,\alpha}.$$

Two-Sample Hotelling's T^2 ($\Sigma_1 = \Sigma_2$)

- The two-sample Hotelling's T^2 statistic can be simplified if we know the two populations share the same variance-covariance matrix, i.e.

$$\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2.$$

- When the sample sizes n_1 and n_2 are not large, the F-statistic will be

$$F = \frac{n_1+n_2-m-1}{m(n_1+n_2-2)} T^2 \sim F_{m,n_1+n_2-m-1}.$$

- In other words, the DOF ν equals $(n_1+n_2 - m - 1)$ when $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2$.

- At a given significance level α, we reject the null hypothesis $H_0\colon \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$ if

$$F > F_{m,n_1+n_2-m-1,\alpha}.$$

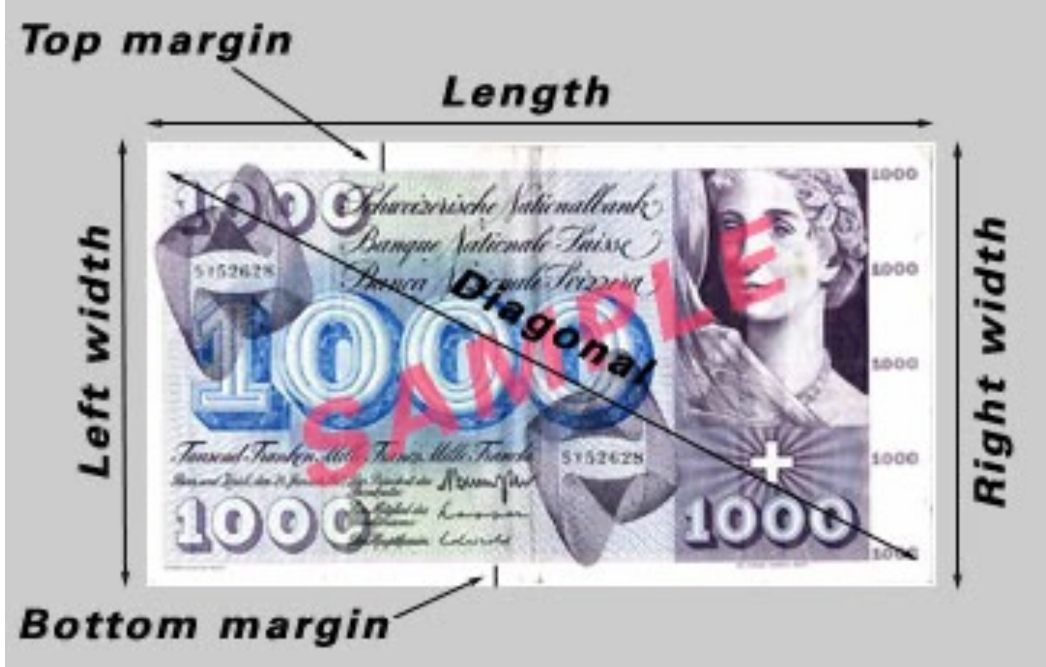
Example: Swiss Bank Notes

- This is a dataset that contains six variables measured on 100 genuine and 100 counterfeit old Swiss 1000-franc bank notes.

Six variables are measured:

1. Length of the note
2. Width of the Left-Hand side of the note
3. Width of the Right-Hand side of the note
4. Width of the Bottom Margin
5. Width of the Top Margin
6. Diagonal Length of Printed Area

Objective: To determine if counterfeit notes can be distinguished from the genuine Swiss bank notes.



Objective: To determine if counterfeit notes can be distinguished from the genuine Swiss bank notes.

- One way to answer the question in our objective is to test if the genuine and counterfeit notes have different population means.

- Denote the population mean vectors of genuine and counterfeit populations as $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$, respectively. Denote $\overline{\boldsymbol{x}}_1$, $\overline{\boldsymbol{x}}_2$, $\mathbf{S}_1$ and $\mathbf{S}_2$ the corresponding sample statistics.

- We are interested in testing the equivalence of the two population means:

$$H_0: \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 \quad \text{vs} \quad H_1: \boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2.$$

- To simplify our calculation, let us first assume the two populations share the same population variance-covariance matrices: (This assumption can be wrong!)

$$\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2.$$

A Peek of the Dataset

- First, let us have a peak at the dataset. BN1 to BN100 are genuine banknotes and BN101 to BN200 are counterfeit banknotes.

	Length	Left	Right	Bottom	Top	Diagonal
BN1	214.8	131	131.1	9	9.7	141
BN2	214.6	129.7	129.7	8.1	9.5	141.7
BN3	214.8	129.7	129.7	8.7	9.6	142.2
BN4	214.8	129.7	129.6	7.5	10.4	142
BN5	215	129.6	129.7	10.4	7.7	141.8

	Length	Left	Right	Bottom	Top	Diagonal
BN101	214.4	130.1	130.3	9.7	11.7	139.8
BN102	214.9	130.5	130.2	11	11.5	139.5
BN103	214.9	130.3	130.1	8.7	11.7	140.2
BN104	215	130.4	130.6	9.9	10.9	140.3
BN105	214.7	130.2	130.3	11.8	10.9	139.7

Testing Procedure

- We calculate the sample means of each variable from the genuine and counterfeit samples, respectively. The results are displayed in the table below.

Variables	Genuine sample mean	Counterfeit sample mean
Length	214.969	214.823
Left Width	129.943	130.300
Right Width	129.720	139.193
Bottom Margin	8.305	10.530
Top Margin	10.168	11.133
Diagonal	141.517	139.450

Testing Procedure (Sample Covariance Matrix)

- We calculate the sample variance covariance matrices from the genuine and counterfeit samples, respectively.

$$S_1 = \begin{pmatrix} 0.150 & 0.058 & 0.057 & 0.057 & 0.014 & 0.005 \\ 0.058 & 0.133 & 0.086 & 0.057 & 0.049 & -0.043 \\ 0.057 & 0.086 & 0.126 & 0.058 & 0.031 & -0.024 \\ 0.057 & 0.057 & 0.058 & 0.413 & -0.263 & -0.000 \\ 0.014 & 0.049 & 0.031 & -0.263 & 0.421 & -0.075 \\ 0.005 & -0.043 & -0.024 & -0.000 & -0.075 & 0.200 \end{pmatrix}$$

$$S_2 = \begin{pmatrix} 0.124 & 0.032 & 0.024 & -0.101 & 0.019 & 0.012 \\ 0.032 & 0.065 & 0.047 & -0.024 & -0.012 & -0.005 \\ 0.024 & 0.047 & 0.089 & -0.019 & 0.000 & 0.034 \\ -0.101 & -0.024 & -0.019 & 1.281 & -0.490 & 0.238 \\ 0.019 & -0.012 & 0.000 & -0.490 & 0.404 & -0.022 \\ 0.012 & -0.005 & 0.034 & 0.238 & -0.022 & 0.311 \end{pmatrix}$$

Genuine Sample

Counterfeit Sample

- Then we can calculate $\mathbf{S}_{\mathrm{T}} = \left\{\frac{1}{n_1}\mathbf{S}_1 + \frac{1}{n_2}\mathbf{S}_2\right\}$.

Conclusion for Swiss Bank Notes Data

- Calculate T^2-statistic

$$T^2 = (\bar{\boldsymbol{x}}_1 - \bar{\boldsymbol{x}}_2)' \left\{\frac{1}{n_1}\mathbf{S}_1 + \frac{1}{n_2}\mathbf{S}_2\right\}^{-1} (\bar{\boldsymbol{x}}_1 - \bar{\boldsymbol{x}}_2) = 2412.45.$$

- Calculate F-statistic

$$F = \frac{n_1+n_2-p-1}{p(n_1+n_2-2)} T^2 = 391.26.$$

- Degree of Freedom

$$m = 6,\ n_1 + n_2 - m - 1 = 193.$$

- Calculate critical value ($\alpha = 0.05$)

$$F_{m,n_1+n_2-m-1,\alpha} = 2.15.$$

- Decision: We reject the null hypothesis $H_0\colon \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$ as

$$F > F_{m,n_1+n_2-m-1,\alpha}.$$

- Conclusion: The counterfeit notes can be distinguished from the genuine notes on at least one of the measurements.