Inference about a Mean Vector

Univariate Case and Multivariate Case

Santiago Alférez

Análisis Estadístico de Datos

MACC
Universidad del Rosario

Agosto de 2020

Outline

- Goal
- Univariate Case
- Multivariate Case
  - Hotelling T^2
  - Likelihood Ratio test
  - Comparison/relationship

Goal

Inference: To make a valid conclusion about the means of a population based on a sample (information about the population).

When we have p correlated variables, they must be analyzed jointly.

Simultaneous analysis yields stronger tests, with better error control.

The tests covered in this set of notes are all of the form:

$$H_o{:}\boldsymbol{\mu} = \boldsymbol{\mu}_o$$

where $\boldsymbol{\mu}_{p\times 1}$ vector of populations means and $\boldsymbol{\mu}_{o,p\times 1}$ is the some specified values under the null hypothesis.

Univariate Case

We're interested in the mean of a population and we have a random sample of n observations from the population,

$$X_1, X_2, \ldots, X_n$$

where (i.e., Assumptions):

- Observations are independent (i.e., X_j is independent from $X_{j'}$ for $j \neq j'$).
- Observations are from the same population; that is,

$$E(X_j) = \mu \text{ for all } j$$

- If the sample size is "small", we'll also assume that

$$X_j \sim \mathcal{N}(\mu, \sigma^2)$$

Hypothesis & Test

- Hypothesis:

$$H_o : \mu = \mu_o \qquad \text{versus} \qquad H_1 : \mu \neq \mu_o$$

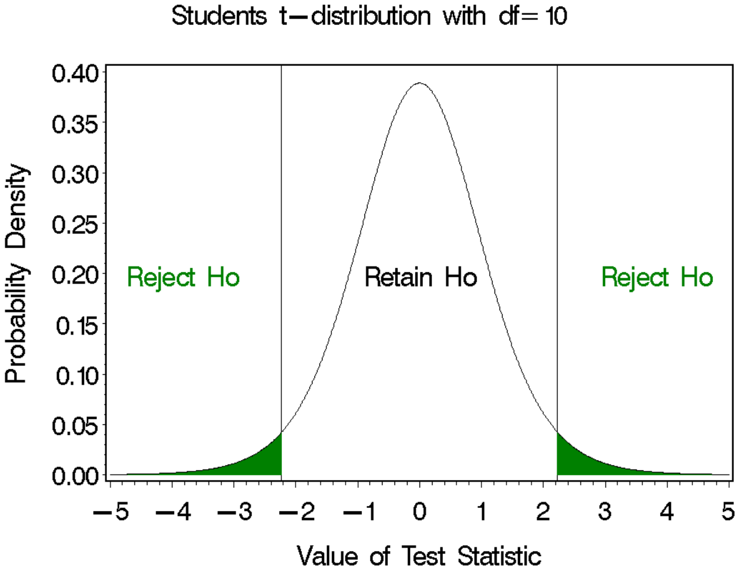
  where μ_o is some specified value. In this case, H_1 is 2–sided alternative.
- Test Statistic:

$$t = \frac{\bar{X} - \mu_o}{s/\sqrt{n}}$$

  where $\bar{X} = (1/n)\sum_{j=1}^{n} X_j$ and
  $s = \sqrt{(1/(n-1))\sum_{j-1}^{n}(X_j - \bar{X})^2}$
- Sampling Distribution: If H_o and assumptions are true, then the sampling distribution of t is Student's - t distribution with $df = n - 1$.
- Decision: Reject H_o when t is "large" (i.e., small p–value).

Picture of Decision

Each green area $= \alpha/2 = .025\ldots$



Confidence Interval

Confidence Interval: A region or range of plausible μ's (given observations/data). The set of all μ's such that

$$\left|\frac{\bar{x}-\mu_o}{s/\sqrt{n}}\right| \leq t_{n-1,(\alpha/2)}$$

where $t_{n-1,(\alpha/2)}$ is the upper $(\alpha/2)100\%$ percentile of Student's t-distribution with $df = n-1$. ...OR

$$\left\{\mu_o \text{ such that } \bar{x}-t_{n-1,(\alpha/2)}\frac{s}{\sqrt{n}} \leq \mu_o \leq \bar{x}+t_{n-1,(\alpha/2)}\frac{s}{\sqrt{n}}\right\}$$

A $100(1-\alpha)^{th}$ confidence interval or region for μ is

$$\left(\bar{x}-t_{n-1,(\alpha/2)}\frac{s}{\sqrt{n}}, \quad \bar{x}+t_{n-1,(\alpha/2)}\frac{s}{\sqrt{n}}\right)$$

Before for sample is selected, the ends of the interval depend on random variables $\bar{X}$'s and s; this is a random interval. $100(1-\alpha)^{th}$ percent of the time such intervals with contain the "true" mean μ

Prepare for Jump to p Dimensions

Square the test statistic t:

$$t^2 = \frac{(\bar{x} - \mu_o)^2}{s^2/n} = n(\bar{x} - \mu_o)(s^2)^{-1}(\bar{x} - \mu_o)$$

So t^2 is a squared statistical distance between the sample mean $\bar{x}$ and the hypothesized value μ_o.

Remember that $t_{df}^2 = \mathcal{F}_{1,df}$?

That is, the sampling distribution of

$$t^2 = n(\bar{x} - \mu_o)(s^2)^{-1}(\bar{x} - \mu_o) \sim \mathcal{F}_{1,n-1}.$$

Multivariate Case: Hotelling's T^2

For the extension from the univariate to multivariate case, replace scalars with vectors and matrices:

$$T^2 = n(\bar{\mathbf{X}} - \boldsymbol{\mu}_o)'\mathbf{S}^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu}_o)$$

- $\bar{\mathbf{X}}_{p\times 1} = (1/n)\sum_{j=1}^{n} \mathbf{X}_j$
- $\boldsymbol{\mu}_{o,(p\times 1)} = (\mu_{1o}, \mu_{2o}, \ldots, \mu_{po})$
- $S_{p\times p} = \frac{1}{n-1}\sum_{j=1}^{n}(\mathbf{X}_j - \bar{\mathbf{X}})(\mathbf{X}_j - \bar{\mathbf{X}})'$

T^2 is "Hotelling's T^2"

The sample distribution of T^2

$$T^2 \sim \frac{(n-1)p}{n-p}\mathcal{F}_{p,(n-p)}$$

We can use this to test $H_o : \boldsymbol{\mu} = \boldsymbol{\mu}_o$... assuming that observations are a random sample from $\mathcal{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ $i.i.d.$

Hotelling's T^2

Since

$$T^2 \sim \frac{(n-1)p}{n-p}\mathcal{F}_{p,(n-p)}$$

We can compute T^2 and compare it to

$$\frac{(n-1)p}{n-p}\mathcal{F}_{p,(n-p)}(\alpha)$$

OR use the fact that

$$\frac{n-p}{(n-1)p}T^2 \sim \mathcal{F}_{p,(n-p)}$$

Compute T^2 as

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

and the

$$p\text{–value } = \text{Prob}\left\{\mathcal{F}_{p,(n-p)} \geq \frac{(n-p)}{(n-1)p}T^2\right\}$$

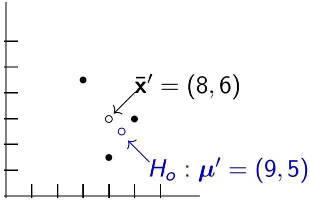
Reject H_o when p-value is small (i.e., when T^2 is large).

Example 1

$n = 3$ and $p = 2$

$$\text{Data: } \mathbf{X} = \begin{pmatrix} 6 & 9 \\ 10 & 6 \\ 8 & 3 \end{pmatrix}$$

$$H_o : \boldsymbol{\mu} = \begin{pmatrix} 9 \\ 5 \end{pmatrix}$$



Assuming data come from a multivariate normal distribution and independent observations,

$$\bar{\mathbf{x}} = \begin{pmatrix} 8 \\ 6 \end{pmatrix} \qquad \mathbf{S} = \begin{pmatrix} 4 & -3 \\ -3 & 9 \end{pmatrix}$$

$$\mathbf{S}^{-1} = \frac{1}{4(9) - (-3)(-3)} \begin{pmatrix} 9 & 3 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 1/3 & 1/9 \\ 1/9 & 4\ 27 \end{pmatrix}$$

Example 1

$$
\begin{aligned}
T^2 &= n(\bar{\mathbf{x}} - \boldsymbol{\mu}_o)'\mathbf{S}^{-1}(\bar{\mathbf{x}} - \boldsymbol{\mu}_o) \\
&= 3\,((8-9),(6-5)) \begin{pmatrix} 1/3 & 1/9 \\ 1/9 & 4/27 \end{pmatrix} \begin{pmatrix} (8-9) \\ (6-5) \end{pmatrix} \\
&= 3(-1,1) \begin{pmatrix} 1/3 & 1/9 \\ 1/9 & 4/27 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \end{pmatrix} \\
&= 3(7/27) = 7/9
\end{aligned}
$$

Value we need for $\alpha = .05$ is $\mathcal{F}_{2,1}(.05) = 199.51$.

$$\frac{(3-1)2}{3-2} 199.51 = 4(199.51) = 798.04.$$

Since $T^2 \sim \frac{(n-1)p}{(n-p)} \mathcal{F}_{p,n-p}$, we can compare our T^2 to 798.04.

Alternatively, we could compute p-value: compare $.25(7/9) = 0.194$ to $\mathcal{F}_{2,1}$ and we get p-value $= .85$.

Do not reject H_o. ($\bar{x}$ and $\boldsymbol{\mu}$ are "close" in the figure).

Example 2: WAIS and n = 101 elderly subjects

There are two variables, verbal and performance scores for $n = 101$ elderly subjects aged 60–64 on the Wechsler Adult Intelligence test (WAIS).

Assume that the data are from a bivariate normal distribution with unknown mean vector $\boldsymbol{\mu}$ and unknown covariance matrix $\boldsymbol{\Sigma}$.

$$H_o : \boldsymbol{\mu} = \begin{pmatrix} 60 \\ 50 \end{pmatrix} \qquad \text{versus} \qquad H_o : \boldsymbol{\mu} \neq \begin{pmatrix} 60 \\ 50 \end{pmatrix}$$

Sample mean vector and covariance matrix:

$$\bar{\mathbf{x}} = \begin{pmatrix} 55.24 \\ 34.97 \end{pmatrix} \qquad \text{and} \qquad \mathbf{S} = \begin{pmatrix} 210.54 & 126.99 \\ 126\ 99 & 119\ 68 \end{pmatrix}$$

Example 2: WAIS and n = 101 elderly subjects

We need

$$\mathbf{S}^{-1} = \begin{pmatrix} .01319 & -.0140 \\ -.0140 & .02321 \end{pmatrix}$$

Compute test statistic:

$$\begin{aligned} T^2 &= n(\bar{\mathbf{x}} - \boldsymbol{\mu})'\mathbf{S}^{-1}(\bar{\mathbf{x}} - \boldsymbol{\mu}) \\ &= 101\,((55.24 - 60), (34.97 - 50))\begin{pmatrix} .01319 & -.0140 \\ -.0140 & .02321 \end{pmatrix}\begin{pmatrix} 55.24 - 60 \\ 34.97 - 50 \end{pmatrix} \\ &= 357.43 \end{aligned}$$

So to test the hypothesis, compute

$$\frac{(n-p)}{(n-1)p} T^2 = \frac{(101-2)}{(101-1)2} 357.43 = 176.93$$

Under the null hypothesis, this is distributed as $\mathcal{F}_{p,(n-p)}$. Since $\mathcal{F}_{2,99}(\alpha = .05) = 3.11$, we reject the null hypothesis.

Example 3: testing a multivariate mean vector with T^2 using the Sweat Data

Perspiration from 20 healthy females was analyzed. Three components, $X_1 =$ sweat rate, $X_2 =$ sodium content, and $X_3 =$ potassium content, were measured, and the results, which we call the sweat data, are presented in the following Table (next slide).

Test the hypothesis $H_0 : \boldsymbol{\mu}' = [4, 50, 10]$ against $H_1 : \boldsymbol{\mu}' \neq [4, 50, 10]$ at level of significance $\alpha = .10$.

Example 3: Sweat Data

Individual	X_1 (Sweat rate)	X_2 (Sodium)	X_3 (Potassium)
1	3.7	48.5	9.3
2	5.7	65.1	8.0
3	3.8	47.2	10.9
4	3.2	53.2	12.0
5	3.1	55.5	9.7
6	4.6	36.1	7.9
7	2.4	24.8	14.0
8	7.2	33.1	7.6
9	6.7	47.4	8.5
10	5.4	54.1	11.3
11	3.9	36.9	12.7
12	4.5	58.8	12.3
13	3.5	27.8	9.8
14	4.5	40.2	8.4
15	1.5	13.5	10.1
16	8.5	56.4	7.1
17	4.5	71.6	8.2
18	6.5	52.8	10.9
19	4.1	44.1	11.2
20	5.5	40.9	9.4

Back to the Univariate Case

Recall that for the univariate case

$$t = \frac{\bar{X} - \mu_o}{s/\sqrt{n}} \quad \text{or} \quad t^2 = \frac{(\bar{X} - \mu_o)^2}{s^2/n} = n(\bar{X} - \mu_o)(s^2)^{-1}(\bar{X} - \mu_o)$$

Since $\bar{X} \sim \mathcal{N}(\mu, (1/n)\sigma^2)$,

$$\sqrt{n}(\bar{X} - \mu_o) \sim \mathcal{N}(\sqrt{n}(\mu - \mu_o), \sigma^2)$$

This is is a linear function of $\bar{X}$, which is a random variable.

We also know that

$$(n-1)s^2 = \sum_{j=1}^{n}(X_j - \bar{X})^2 \sim \sigma^2\chi^2_{(n-1)}$$

because

$$\frac{\sum_{j=1}^{n}(X_j - \bar{X})^2}{\sigma^2} = \sum_{j=1}^{n} Z_j^2 \sim \chi^2_{(n-1)}$$

Back to the Univariate Case continued

So

$$s^2 = \frac{\sum_{j=1}^{n}(X_j - \bar{X})^2}{n-1} = \frac{\text{chi-square random variable}}{\text{degrees of freedom}}$$

Putting this all together, we find

$$t^2 = \left(\begin{array}{l}\text{normal}\\ \text{random}\\ \text{variable}\end{array}\right)\left(\frac{\text{chi-square random varible}}{\text{degress of freedom}}\right)^{-1}\left(\begin{array}{l}\text{normal}\\ \text{random}\\ \text{variable}\end{array}\right)$$

Now we'll go through the same thing but with the multivariate case...

The Multivariate Case

$$T^2 = \sqrt{n}(\bar{\mathbf{X}} - \boldsymbol{\mu}_o)'(\mathbf{S})^{-1}\sqrt{n}(\bar{\mathbf{X}} - \boldsymbol{\mu}_o)$$

Since $\bar{\mathbf{X}} \sim \mathcal{N}_p(\boldsymbol{\mu}, (1/n)\boldsymbol{\Sigma})$ and $\sqrt{n}(\bar{\mathbf{X}} - \boldsymbol{\mu}_o)$ is a linear combination of $\bar{\mathbf{X}}$,

$$\sqrt{n}(\bar{\mathbf{X}} - \boldsymbol{\mu}_o) \sim \mathcal{N}_p(\sqrt{n}(\boldsymbol{\mu} - \boldsymbol{\mu}_o), \boldsymbol{\Sigma})$$

Also

$$\begin{aligned}
\mathbf{S} &= \frac{\sum_{j=1}^{n}(\mathbf{X}_j - \bar{\mathbf{X}})(\mathbf{X}_j - \bar{\mathbf{X}})'}{(n-1)} \\
&= \frac{\sum_{j=1}^{n}\mathbf{Z}_j\mathbf{Z}_j'}{(n-1)} \\
&= \left(\frac{\text{Wishart random matrix with df} = n-1}{\text{degrees of freedom}} \right)
\end{aligned}$$

where $\mathbf{Z}_j \sim \mathcal{N}_p(\mathbf{0}, \boldsymbol{\Sigma})$ *i.i.d.*... if H_o is true.

The Multivariate Case continued

Recall that a Wishart distribution is a matrix generalization of the chi-square distribution.

The sampling distribution of $(n-1)\mathbf{S}$ is Wishart where

$$\mathbf{W}_m(\cdot|\mathbf{\Sigma}) = \sum_{j=1}^{m} \mathbf{Z}_j \mathbf{Z}_j'$$

where $\mathbf{Z}_j \sim \mathcal{N}_p(\mathbf{0}, \mathbf{\Sigma})$ *i.i.d.*.
So,

$$T^2 = \begin{pmatrix} \text{multiavirate} \\ \text{normal} \\ \text{random} \\ \text{vector} \end{pmatrix} \left(\frac{\text{Wishart random matrix}}{\text{degress of freedom}} \right)^{-1} \begin{pmatrix} \text{multiavirate} \\ \text{normal} \\ \text{random} \\ \text{vector} \end{pmatrix}$$

Invariance of T^2

T^2 is invariant with respect to change of location (i.e., mean) or scale (i.e. covariance matrix); that is, a T^2 is invariant by linear transformation.

Rather than $\mathbf{X}_{p\times 1}$, we may want to consider

$$\mathbf{Y}_{p\times 1} = \underbrace{\mathbf{C}_{p\times p}}_{\textit{scale}}\mathbf{X}_{p\times 1} + \underbrace{\mathbf{d}_{p\times 1}}_{\textit{location}}$$

where $\mathbf{C}$ is non-singular (or equivalently $|\mathbf{C}| > 0$, or $\mathbf{C}$ has p linearly independent rows (columns), or $\mathbf{C}^{-1}$ exists).

$$v\mu_y = \mathbf{C}\boldsymbol{\mu}_x + \mathbf{d} \qquad \text{and} \qquad \boldsymbol{\Sigma}_y = \mathbf{C}\boldsymbol{\Sigma}_x\mathbf{C}'$$

The T^2 for the Y–data is exactly the same as the T^2 for the X–data (see text for proof).

This result it true for the univariate t-test.

Likelihood Ratio

- Another approach to testing null hypothesis about mean vector $\boldsymbol{\mu}$ (as well as other multivariate tests in general).
- It's equivalent to Hotelling's T^2 for $H_o : \boldsymbol{\mu} = \boldsymbol{\mu}_o$ or $H_o : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$.
- It's more general than T^2 in that it can be used to test other hypotheses (e.g., those regarding $\boldsymbol{\Sigma}$) and in different circumstances.
- Foreshadow: When testing more than 1 or 2 mean vectors, there are lots of different test statistics (about 5 common ones).
- T^2 and likelihood ratio tests are based on different underlying principles.

Underlying Principles

T^2 is based on the <u>union-intersection</u> principle, which takes a multivariate hypothesis and turns it into a univariate problem by considering linear combinations of variables. i.e.,

$$T^2 = \mathbf{a}'(\bar{\mathbf{X}} - \boldsymbol{\mu}_o)$$

is a linear combination.

We select the combination vector **a** that lead to the largest possible value of T^2. (We'll talk more about this later). The emphasis is on the "direction of maximal difference".

The likelihood ratio test the emphasis is on overall difference.

Plan: First talk about the basic idea behind Likelihood ratio tests and then we'll apply it to the specific problem of testing $\boldsymbol{\mu} = \boldsymbol{\mu}_o$.

Basic idea of Likelihood Ratio Tests

- Θ_o = a set of unknown parameters under H_o (e.g., $\boldsymbol{\Sigma}$).
- Θ = the set of unknown parameters under the alternative hypothesis (model), which is more general (e.g., $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$).
- $\mathcal{L}(\cdot)$ is the likelihood function. It is a function of parameters that indicates "how likely Θ (or Θ_o) is given the data".
- $\mathcal{L}(\Theta) \geq \mathcal{L}(\Theta_o)$.
  - The more general model/hypothesis is always more (or equally) likely than the more restrictive model/hypothesis.

The Likelihood Ratio Statistic is

$$\Lambda = \frac{\max \mathcal{L}(\Theta_o)}{\max \mathcal{L}(\Theta)} \quad \rightarrow \quad \begin{array}{ll} \bar{\mathbf{X}} = \hat{\boldsymbol{\mu}} & \text{MLE of mean} \\ \mathbf{S}_n = \hat{\boldsymbol{\Sigma}} & \text{MLE of covariance matrix} \end{array}$$

If Λ is "small", then the data are not likely to have occurred under $H_o \longrightarrow$ Reject H_o.

If Λ is "large", then the data are likely to have occurred under $H_o \longrightarrow$ Retain H_o.

Likelihood Ratio Test for Mean Vector

Let $\mathbf{X}_j \sim \mathcal{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and *i.i.d.*

$$\Lambda = \frac{\max_{\boldsymbol{\Sigma}}[\mathcal{L}(\boldsymbol{\mu}_o, \boldsymbol{\Sigma})]}{\max_{\boldsymbol{\mu},\boldsymbol{\Sigma}}[\mathcal{L}(\boldsymbol{\mu}, \boldsymbol{\Sigma})]}$$

where

- $\max_{\boldsymbol{\Sigma}} =$ the maximum of $\mathcal{L}(\cdot)$ over all possible $\boldsymbol{\Sigma}$'s.
- $\max_{\boldsymbol{\mu},\boldsymbol{\Sigma}} =$ the maximum of $\mathcal{L}(\cdot)$ over all possible $\boldsymbol{\mu}$'s & $\boldsymbol{\Sigma}$'s.

$$\Lambda = \left(\frac{|\hat{\boldsymbol{\Sigma}}|}{|\hat{\boldsymbol{\Sigma}}_o|}\right)^{n/2}$$

where

- $\hat{\boldsymbol{\Sigma}} =$ MLE of $\boldsymbol{\Sigma} = (1/n)\sum_{j=1}^n(\mathbf{X}_j - \bar{\mathbf{X}})(\mathbf{X}_j - \bar{\mathbf{X}})' = \mathbf{S}_n$
- $\hat{\boldsymbol{\Sigma}}_o =$ MLE of $\boldsymbol{\Sigma}$ assuming that $\boldsymbol{\mu} = \boldsymbol{\mu}_o$
  $= (1/n)\sum_{j=1}^n(\mathbf{X}_j - \boldsymbol{\mu}_o)(\mathbf{X}_j - \boldsymbol{\mu}_o)'$

Likelihood Ratio Test for Mean Vector

$$\Lambda = \left(\frac{|\hat{\mathbf{\Sigma}}|}{|\hat{\mathbf{\Sigma}}_o|} \right)^{n/2}$$

$$\Lambda = (\text{ratio of two generalized sample variances})^{n/2}$$

- If $\boldsymbol{\mu}_o$ is really "far" from $\boldsymbol{\mu}$, then $|\hat{\mathbf{\Sigma}}_o|$ will be much larger than $|\hat{\mathbf{\Sigma}}|$, which uses a "good" estimator of $\boldsymbol{\mu}$ (i.e., $\bar{\mathbf{X}}$).
- The likelihood ratio statistic Λ is called "Wilk's Lambda" for the special case of testing hypotheses about mean vectors.
- For large samples (i.e., large n),

$$-2\ln(\Lambda) \sim \chi^2_p,$$

which can be used to test $H_o : \boldsymbol{\mu} = \boldsymbol{\mu}_o$

Degrees of Freedom for LR Test

We need to consider the number of parameter estimates under each hypothesis:
The alternative hypothesis ("full model") ,

$$\Theta = \{\boldsymbol{\mu}, \boldsymbol{\Sigma}\} \longrightarrow p \text{ means} \quad + \quad \frac{p(p-1)}{2} \text{ covariances}$$

The null hypothesis,

$$\Theta_o = \{\boldsymbol{\Sigma}\} \longrightarrow \frac{p(p-1)}{2} \text{ covariances}$$

$$\begin{aligned} \text{degrees of freedom} = df \quad &= \quad \text{difference between number of parameters} \\ &\quad \quad \text{estimated under each hypothesis} \\ &= \quad p \end{aligned}$$

If the H_o is true and all assumptions valid, then for large samples, $-2\ln(\Lambda) \sim \chi^2_p$.

Example 4: four psychological tests

$n = 64,\ p = 4,\ \bar{\mathbf{x}}' = (14.15, 14.91, 21.92, 22.34)$,

$$\mathbf{S} = \begin{pmatrix} 10.388 & 7.793 & 15.298 & 5.3740 \\ 7.793 & 16.658 & 13.707 & 6.1756 \\ 15.298 & 13.707 & 57.058 & 15.932 \\ 5.374 & 6.176 & 15.932 & 22.134 \end{pmatrix} \quad \& \ \det(\mathbf{S}) = 61952.085$$

Test: $H_o : \boldsymbol{\mu}' = (20, 20, 20, 20)$ versus $H_o : \boldsymbol{\mu}' \neq (20, 20, 20, 20)$

$$\boldsymbol{\Sigma}_o = \frac{1}{n}(\mathbf{X} - \mathbf{1}\boldsymbol{\mu}_o')'(\mathbf{X} - \mathbf{1}\boldsymbol{\mu}_o') = \begin{pmatrix} 44.375 & 37.438 & 3.828 & -8.406 \\ 37.438 & 42.344 & 3.703 & -5.859 \\ 3.828 & 3.703 & 59.859 & 20.187 \\ -8.406 & -5.859 & 20.187 & 27.281 \end{pmatrix}$$

$\det(\boldsymbol{\Sigma}_o) = 518123.8$.

Wilk's Lambda is $\Lambda = (61952.085/518123.8)^{64/2} = 3.047E - 30$, and Comparing $-2\ln(\Lambda) = 135.92659$ to a χ^2_4 gives p–value $<< .01$.

Comparison of T^2 & Likelihood Ratio

Hotelling's T^2 and Wilk's Lambda are functionally related.

Let $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ be a random sample from a $\mathcal{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ population, then the test of $H_o : \boldsymbol{\mu} = \boldsymbol{\mu}_o$ versus $H_A : \boldsymbol{\mu} \neq \boldsymbol{\mu}_o$ based on T^2 is equivalent to the test based on Λ.
The relationship is given by

$$(\Lambda)^{2/n} = \left(1 + \frac{T^2}{(n-1)}\right)^{-1}$$

So,

$$\Lambda = \left(1 + \frac{T^2}{(n-1)}\right)^{-n/2} \quad \text{and} \quad T^2 = (n-1)\Lambda^{-2/n} - (n-1)$$

Since they are inversely related,

- We reject H_o for "large" T^2
- We reject H_o for "small" Λ.

Example 4.1: Comparison of T^2 & Likelihood Ratio

Using our 4 psychological test data, we found that

$$(\Lambda) = 3.047E - 30$$

If we compute Hotelling's T^2 for these data we find that

$$T^2 = 463.88783$$

$$\Lambda = \left(1 + \frac{463.88783}{(64-1)}\right)^{-64/2} = 3.047E - 30$$

and

$$T^2 = (64-1)(3.047E - 30)^{-2/64} - (64-1)$$