One-Sample Hotelling's T^2 Test

ST 558: Multivariate Analytics

Module 3

Lecture 3

Hotelling's T^2 Test

Instead of testing each element of the mean vector separately and then combining these multiple test results, we would like to test the entire vector at once.

Hotelling's T^2 test is a multivariate analogue of the univariate t-test.

- Univariate t-test statistic to test $H_0: \mu = \mu_0$

$$t(\mu_0) = \frac{\sqrt{n}(\bar{X} - \mu_0)}{s} \Rightarrow t^2(\mu_0) = n(\bar{X} - \mu_0)(s^2)^{-1}(\bar{X} - \mu_0)$$

- Multivariate T^2-test statistic to test $H_0: \boldsymbol{\mu} = \boldsymbol{\mu}_0$

$$T^2(\boldsymbol{\mu}_0) = n(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)^T(\mathbf{S})^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)$$

Hotelling's T^2 Test

Instead of testing each element of the mean vector separately and then combining these multiple test results, we would like to test the entire vector at once.

Hotelling's T^2 test is a multivariate analogue of the univariate t-test.

- Univariate t-test statistic to test $H_0: \mu = \mu_0$

$$t(\mu_0) = \frac{\sqrt{n}(\bar{X} - \mu_0)}{s} \Rightarrow t^2(\mu_0) = n(\bar{X} - \mu_0)(s^2)^{-1}(\bar{X} - \mu_0)$$

- Multivariate T^2-test statistic to test $H_0: \boldsymbol{\mu} = \boldsymbol{\mu}_0$

$$T^2(\boldsymbol{\mu}_0) = n(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)^T(\mathbf{S})^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)$$

Hotelling's T^2 Test

Instead of testing each element of the mean vector separately and then combining these multiple test results, we would like to test the entire vector at once.

Hotelling's T^2 test is a multivariate analogue of the univariate t-test.

- Univariate t-test statistic to test $H_0: \mu = \mu_0$

$\bar{\mathbf{X}}$ = Sample mean vector

$t($ … $\Rightarrow t^2(\mu_0) = n(\bar{X} -$ …

$\mathbf{S}$ = Sample covariance matrix

- Multivariate T^2-test statistic to test $H_0: \boldsymbol{\mu} = \boldsymbol{\mu}_0$

$$T^2(\boldsymbol{\mu}_0) = n(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)^T (\mathbf{S})^{-1} (\bar{\mathbf{X}} - \boldsymbol{\mu}_0)$$

Hotelling's T^2 Test

Instead of testing each element of the mean vector separately and then combining these multiple test results, we would like to test the entire vector at once.

Hotelling's T^2 test is a multivariate analogue of the univariate t-test.

- Univariate t-test statistic to test $H_0: \mu = \mu_0$

$$t(\mu_0) = \frac{\sqrt{n}(\bar{X} - \mu_0)}{s} \Rightarrow t^2(\mu_0) = n(\bar{X} - \mu_0)(s^2)^{-1}(\bar{X} - \mu_0)$$

- Multivariate T^2-test statistic to test $H_0: \boldsymbol{\mu} = \boldsymbol{\mu}_0$

$$T^2(\boldsymbol{\mu}_0) = n\underbrace{(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)^T(\mathbf{S})^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)}_{\text{Matrix multiplication}}$$

Hotelling's T^2 Test

Instead of testing each element of the mean vector separately and then combining these multiple test results, we would like to test the entire vector at once.

Hotelling's T^2 test is a multivariate analogue of the univariate t-test.

- Univariate t-test statistic to test $H_0: \mu = \mu_0$

$$t(\mu_0) = \frac{\sqrt{n}(\bar{X} - \mu_0)}{s} \Rightarrow t^2(\mu_0) = n(\bar{X} - \mu_0)(s^2)^{-1}(\bar{X} - \mu_0)$$

- Multivariate T^2-test statistic to test $H_0: \boldsymbol{\mu} = \boldsymbol{\mu}_0$

$$T^2(\boldsymbol{\mu}_0) = n\underbrace{(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)^T}_{(1\times p)}\underbrace{(\mathbf{S})^{-1}}_{(p\times p)}\underbrace{(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)}_{(p\times 1)}$$

Hotelling's T^2 Test

When the null hypothesis is true (that is, under H_0), the statistic

$$T^2(\boldsymbol{\mu}_0) = n(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)^T(\mathbf{S})^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)$$

has (approximately, depending on the underlying population distribution) a scaled F distribution:

$$\frac{n-p}{(n-1)p}T^2(\boldsymbol{\mu}_0) \sim F_{(p,n-p)}$$

Large values of $T^2(\boldsymbol{\mu}_0)$ indicate that the hypothesized mean vector $\boldsymbol{\mu}_0$ is relatively *far* from the observed sample mean $\bar{\mathbf{X}}$, and therefore is less plausible.

Hotelling's T^2 Test

When the null hypothesis is true (that is, under H_0), the statistic

$$T^2(\boldsymbol{\mu}_0) = n(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)^T(\mathbf{S})^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)$$

has (approximately, depending on the underlying population distribution) a scaled F distribution:

$$\frac{n-p}{(n-1)p}T^2(\boldsymbol{\mu}_0) \sim F_{(p,n-p)}$$

F distribution with p numerator degrees of freedom and $n - p$ denominator degrees of freedom

Large values of $T^2(\boldsymbol{\mu}_0)$ indicate that the hypothesized mean vector $\boldsymbol{\mu}_0$ is relatively *far* from the observed sample mean $\bar{\mathbf{X}}$, and therefore is less plausible.

Hotelling's T^2 Test

When the null hypothesis is true (that is, under H_0), the statistic

$$T^2(\boldsymbol{\mu}_0) = n(\overline{\mathbf{X}} - \boldsymbol{\mu}_0)^T(\mathbf{S})^{-1}(\overline{\mathbf{X}} - \boldsymbol{\mu}_0)$$

has (approximately, depending on the underlying population distribution) a scaled F distribution:

$$\frac{n-p}{(n-1)p}T^2(\boldsymbol{\mu}_0) \sim F_{(p,n-p)}$$

Therefore, we ***reject*** the null hypothesis when

$$\frac{n-p}{(n-1)p}T^2(\boldsymbol{\mu}_0) > F_{(p,n-p)}(\alpha)$$

Hotelling's T^2 Test

When the null hypothesis is true (that is, under H_0), the statistic

$$T^2(\boldsymbol{\mu}_0) = n(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)^T(\mathbf{S})^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)$$

has (approximately, depending on the underlying population distribution) a scaled F distribution:

$$\frac{n-p}{(n-1)p}T^2(\boldsymbol{\mu}_0) \sim F_{(p,n-p)}$$

Therefore, we ***reject*** the null hypothesis when

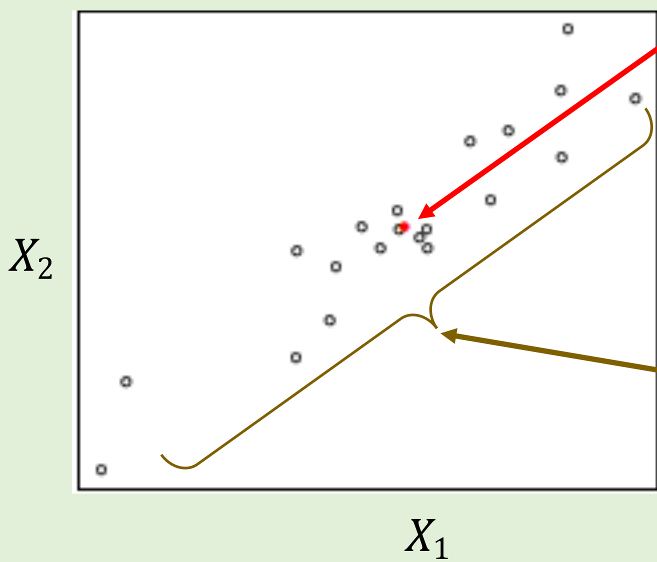
$$\frac{n-p}{(n-1)p}T^2(\boldsymbol{\mu}_0) > F_{(p,n-p)}(\alpha)$$

Upper α quantile of the reference F distribution

$F_{(p,n-p)}(\alpha)$ = `qf(1-alpha, p, n-p)`

Multivariate Hypothesis Testing: Example

Returning to our bivariate example:

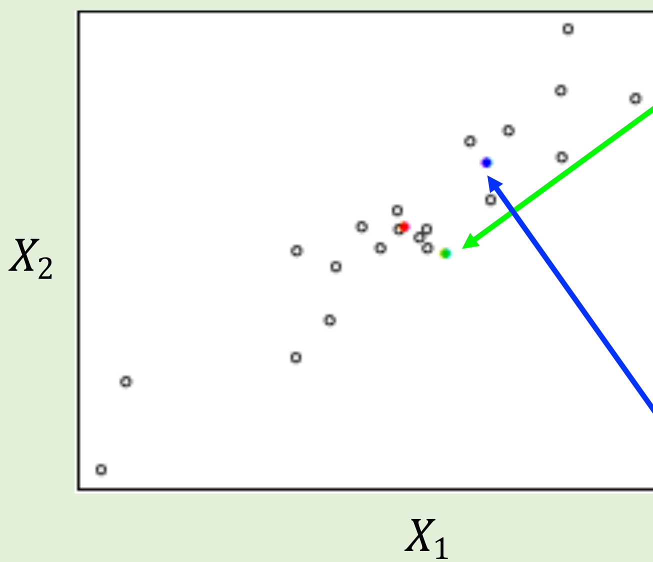


Red dot is the sample mean vector $\bar{\mathbf{X}} = [-0.297, 0.241]^T$

Sample covariance matrix is $\mathbf{S} = \begin{bmatrix} 1.006 & 0.869 \\ 0.869 & 0.862 \end{bmatrix}$

Multivariate Hypothesis Testing: Example

Which of these two points (blue or green) do you think is more plausible as the population mean vector?



Green dot is the first hypothesized mean vector
$\boldsymbol{\mu}_0^{(1)} = [0, 0]^T$

Blue dot is the second hypothesized mean vector
$\boldsymbol{\mu}_0^{(2)} = [0.3, 0.8]^T$

Multivariate Hypothesis Testing: Example

Which of these two points (blue or green) do you think is more plausible as the population mean vector?

Green: $H_0^{(1)}: \boldsymbol{\mu} = \boldsymbol{\mu}_0^{(1)} = [0,0]^T$

$$\frac{n-p}{(n-1)p} T^2\left(\boldsymbol{\mu}_0^{(1)}\right) = 21.867$$

Blue: $H_0^{(2)}: \boldsymbol{\mu} = \boldsymbol{\mu}_0^{(2)} = [0.3, 0.8]^T$

$$\frac{n-p}{(n-1)p} T^2\left(\boldsymbol{\mu}_0^{(2)}\right) = 3.517$$

The critical value for a level $\alpha = 0.05$ test is

$F_{(p,n-p)}(\alpha) = F_{(2,18)}(0.05) =$ `qf(1-0.05, 2, 18)` $= 3.555$

Multivariate Hypothesis Testing: Example

Which of these two points (blue or green) do you think is more plausible as the population mean vector?

Green: $H_0^{(1)}: \boldsymbol{\mu} = \boldsymbol{\mu}_0^{(1)} = [0, 0]^T$

$$\frac{n-p}{(n-1)p} T^2\left(\boldsymbol{\mu}_0^{(1)}\right) = 21.867$$

Reject this null hypothesis at level $\alpha = 0.05$ since $21.867 > 3.555$

Blue: $H_0^{(2)}: \boldsymbol{\mu} = \boldsymbol{\mu}_0^{(2)} = [0.3, 0.8]^T$

$$\frac{n-p}{(n-1)p} T^2\left(\boldsymbol{\mu}_0^{(2)}\right) = 3.517$$

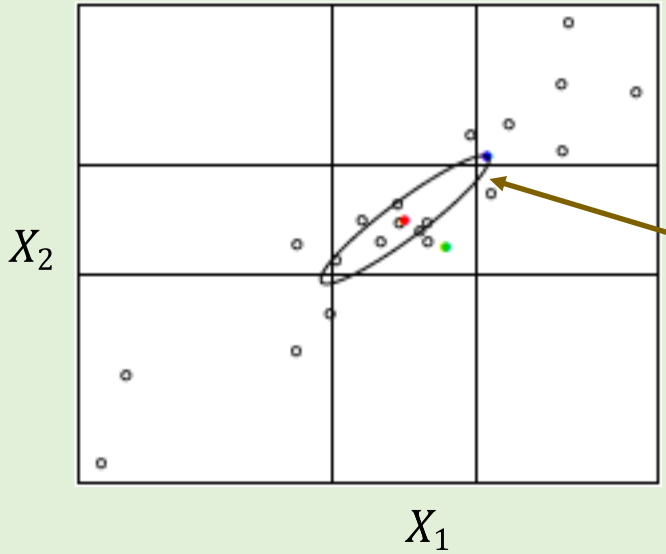
Fail to reject this null hypothesis at level $\alpha = 0.05$ since $3.517 < 3.555$

The critical value for a level $\alpha = 0.05$ test is

$F_{(p,n-p)}(\alpha) = F_{(2,18)}(0.05) =$ `qf(1-0.05, 2, 18)` $= 3.555$

Multivariate Hypothesis Testing: Example

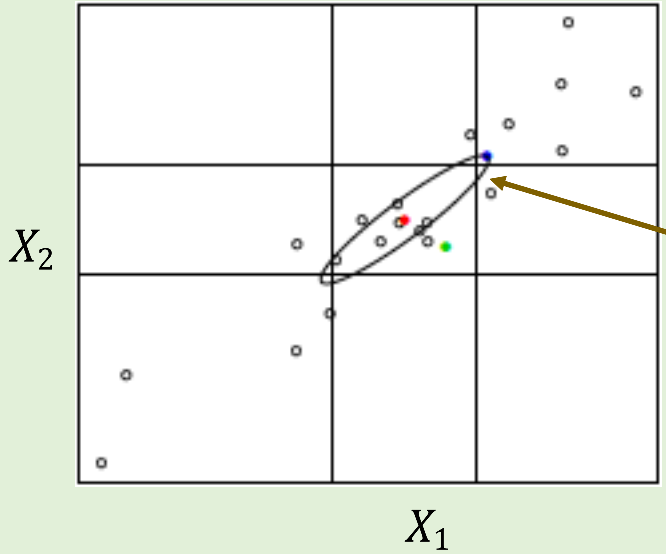
Which of these two points (blue or green) do you think is more plausible as the population mean vector?



Every hypothesis ***inside*** this ellipse would produce a test statistic ***less than*** the critical value, and therefore would **NOT** be rejected

Multivariate Hypothesis Testing: Example

Which of these two points (blue or green) do you think is more plausible as the population mean vector?



Every hypothesis ***outside*** this ellipse would produce a test statistic ***greater than*** the critical value, and therefore **WOULD** be rejected

Multivariate Confidence Regions

Recall the duality between hypothesis tests and confidence regions:

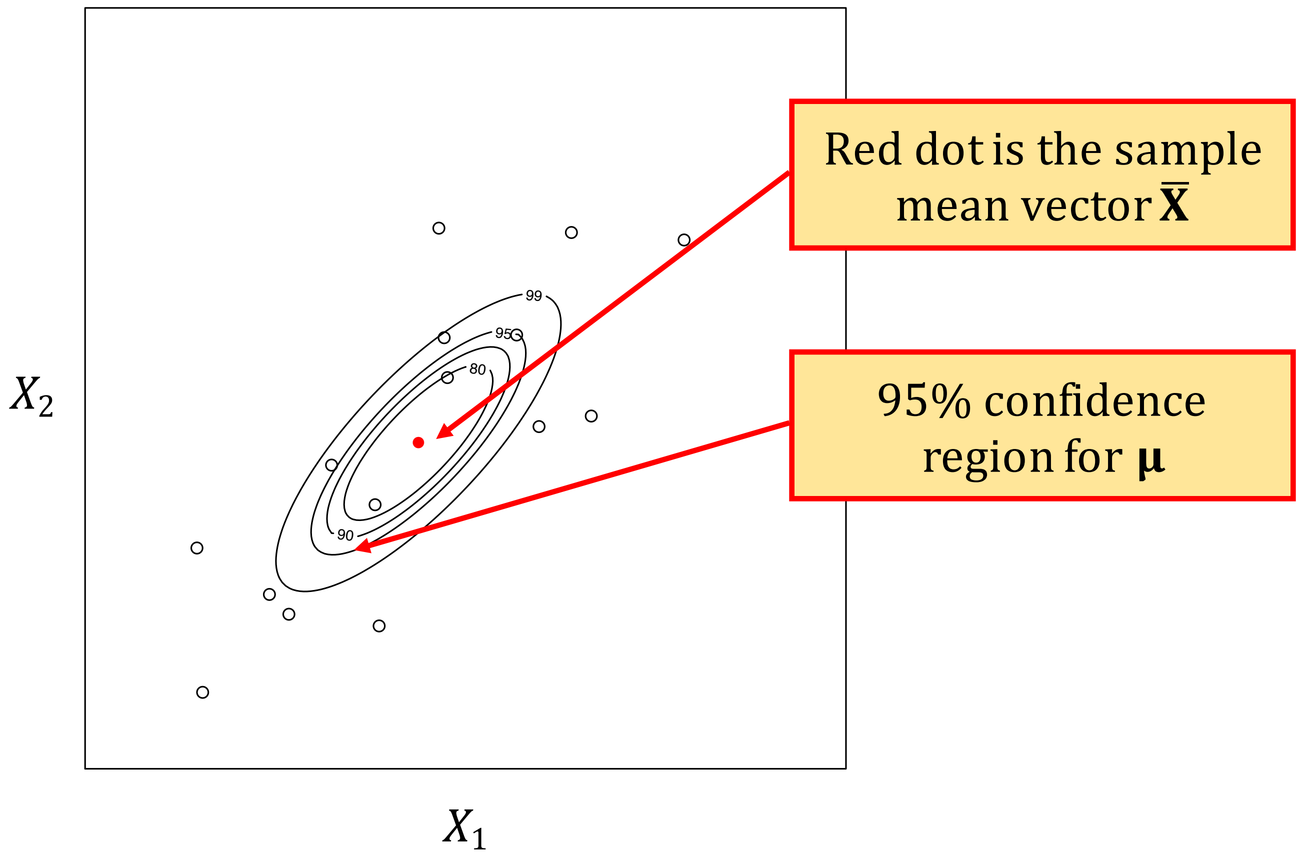
Definition: A $(1-\alpha)100\%$ confidence region for a parameter θ is the set of all values θ_0 for which a test of $H_0\colon \theta = \theta_0$ would not reject at level α.

Therefore, a confidence region for the population mean vector $\boldsymbol{\mu}$ is given by:

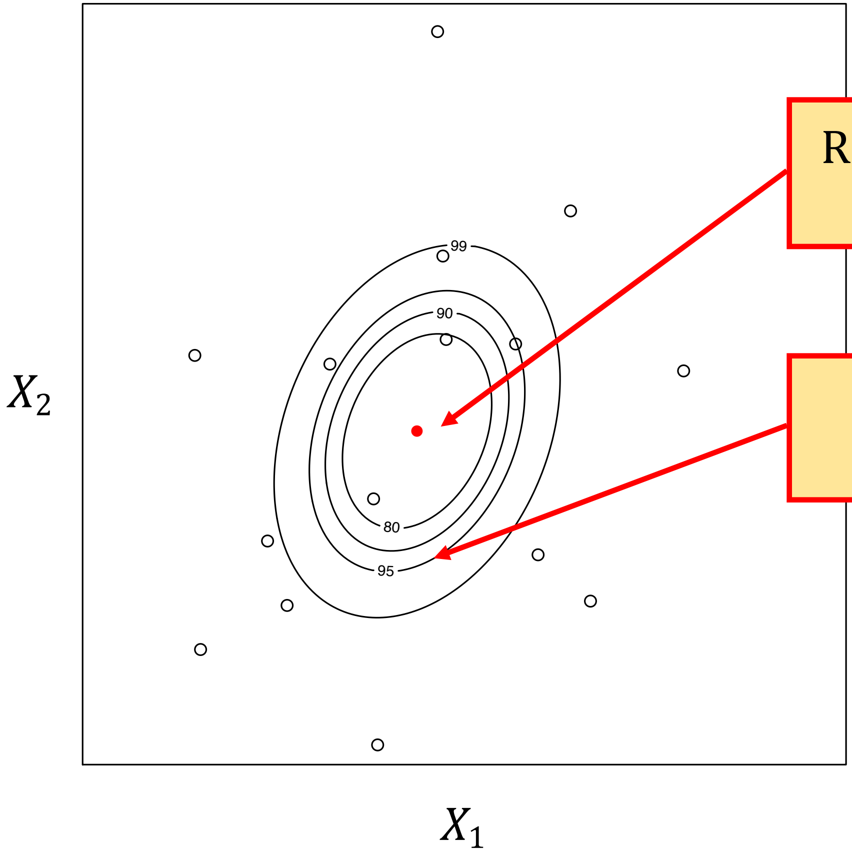
$$C = \left\{ \boldsymbol{\mu}_0 : \frac{n-p}{(n-1)p} T^2(\boldsymbol{\mu}_0) \leq F_{(p,n-p)}(\alpha) \right\}$$

This region will be a p-dimensional ellipsoid centered around the sample mean vector $\overline{\mathbf{X}}$.

Multivariate Confidence Regions



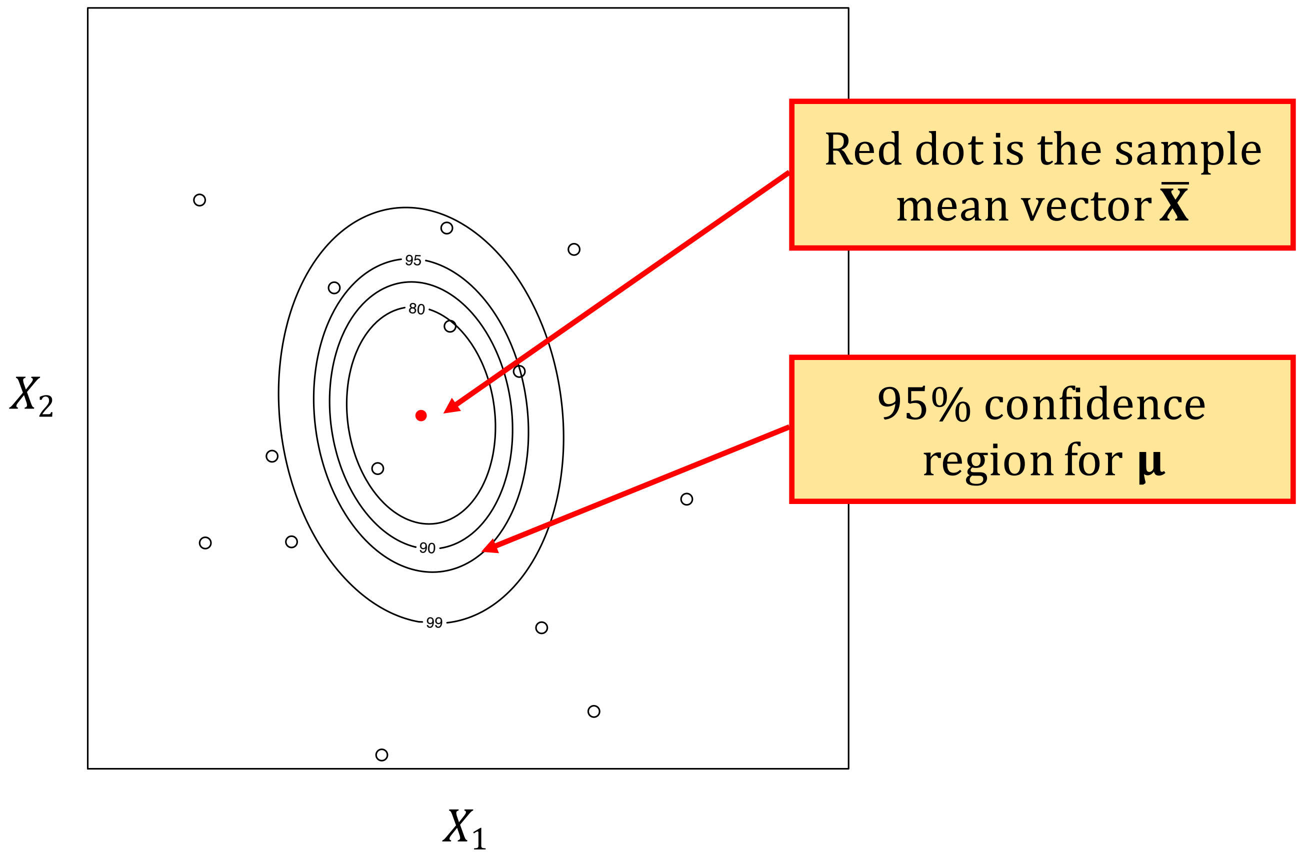
Multivariate Confidence Regions



Red dot is the sample mean vector $\bar{\mathbf{X}}$

95% confidence region for $\boldsymbol{\mu}$

Multivariate Confidence Regions



p-values

Recall that p-values indicate the probability of obtaining a result at least as *extreme* (stronger evidence for the alternative) as the observed outcome *if the null hypothesis is true.*

We can report p-values as an alternative way of giving the result of a test procedure—p-values contain more information than a simple reject/fail to reject decision.

For Hotelling's T^2 test, p-values are obtained by comparing the test statistic to the reference $F_{(p,n-p)}$ distribution:

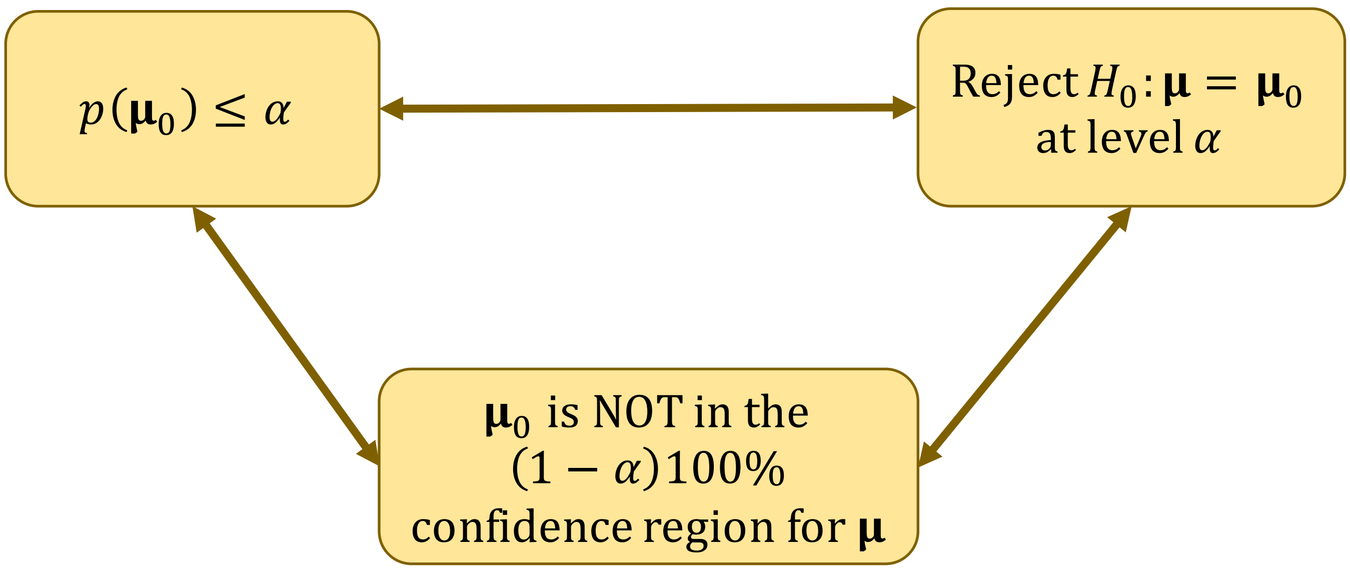
$$p(\boldsymbol{\mu}_0) = P\left(F > \frac{n-p}{(n-1)p} T^2(\boldsymbol{\mu}_0)\right)$$

= `1 – pf((n-p)/((n-1)p)T, p, n-p)`

for $F \sim F_{(p,n-p)}$.

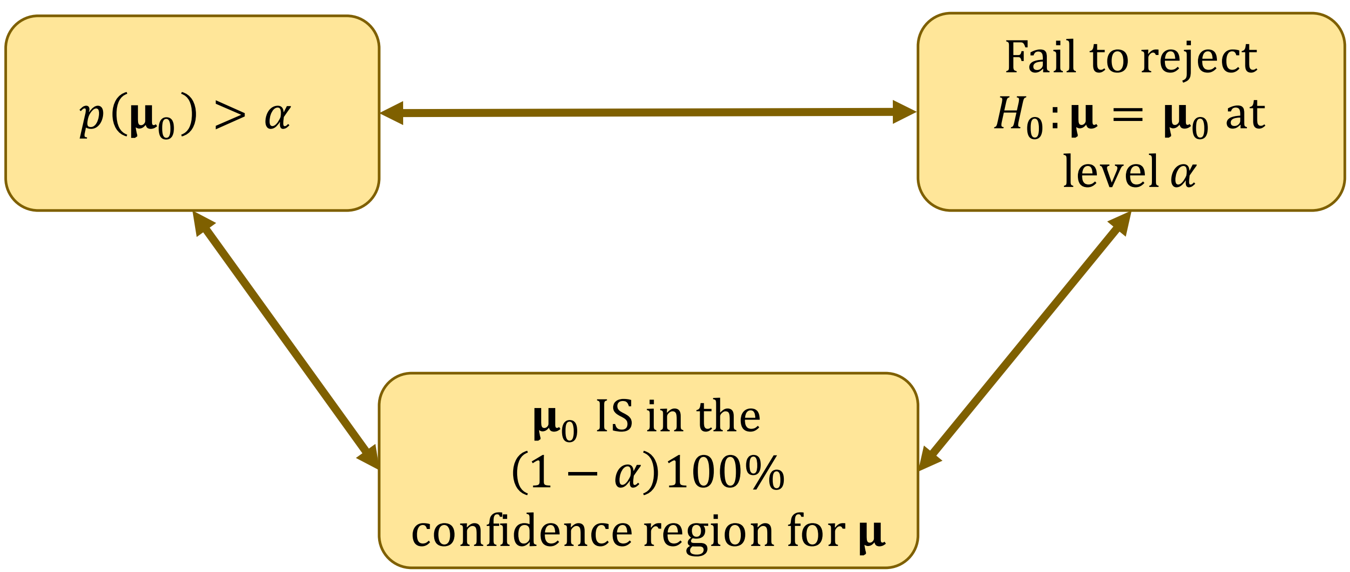
p-values

Reminder: The relationships between p-values and hypothesis test decisions and confidence intervals are as shown below.



p-values

Reminder: The relationships between p-values and hypothesis test decisions and confidence intervals are as shown below.



Performance of Hotelling's T^2 Test

- Hotelling's T^2 test is *exact* if the underlying joint population distribution is exactly multivariate normal
  - p-values are *exact*
  - Type I error rate is *exactly* target level α
  - Confidence region coverage is *exactly* target level $(1-\alpha)100\%$.

- Hotelling's T^2 test is *approximate* if the underlying joint population distribution is any other distribution
  - p-values are *approximate*
  - Type I error rate is *approximately* target level α
  - Confidence region coverage is *approximately* target level $(1-\alpha)100\%$.
  - **This approximation becomes closer and closer to exact as the sample size *n* gets larger.**