# One-Sample Hypothesis Tests for Multivariate Mean Vectors

ST 558: Multivariate Analytics

Module 3

Lecture 2

# Review of Hypothesis Testing Ideas/Vocabulary

We will first quickly review the univariate case to recall the vocabulary and philosophy of hypothesis testing:

- Let $X_1, X_2, \ldots, X_n$ be an iid sample from a population with population mean $\mu$ and population variance $\sigma^2$.
- Question/Goal: Test hypotheses regarding the value of the (unknown) mean parameter $\mu$.

# Review of Hypothesis Testing Ideas/Vocabulary

Hypotheses to be tested:

- *Null hypothesis*:

$$H_0: \mu = \mu_0$$

  for some specified value $\mu_0$

- We compare this null hypothesis to an *Alternative hypothesis*. We can consider several forms for the alternative hypothesis:

$$H_A: \mu > \mu_0 \quad \text{OR}$$
$$H_A: \mu < \mu_0 \quad \text{OR}$$
$$H_A: \mu \neq \mu_0$$

# Review of Hypothesis Testing Ideas/Vocabulary

Hypotheses to be tested:

- *Null hypothesis*:
$$H_0\!: \mu = \mu_0$$
for some specified value $\mu_0$
- We compare this null hypothesis to an *Alternative hypothesis*. We can consider several forms for the alternative hypothesis:
$$H_A\!: \mu > \mu_0 \quad \text{OR}$$
$$H_A\!: \mu < \mu_0 \quad \text{OR}$$
$$\boldsymbol{H_A\!: \mu \neq \mu_0}$$
We will focus on the third of these possible alternatives, the *two-sided alternative hypothesis.*

# Univariate Hypothesis Testing Example

**Example:** We have measured body temperature on $n$ = 12 healthy subjects.

| Subject | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Body Temp. | 94.3 | 95.2 | 98.2 | 96.8 | 99.3 | 98.5 | 99.4 | 97.6 | 99.2 | 98.3 | 97.9 | 98.0 |

We would like to use this data to perform test of the null hypothesis that the population mean (average) body temperature is 98.6° F.

$$H_0: \mu = 98.6$$

$$H_A: \mu \neq 98.6$$

# Review of Hypothesis Testing Ideas/Vocabulary

The result of a hypothesis test is one of two possible decisions/outcomes:

- Reject the null hypothesis $H_0$: conclude that the data do not support the null hypothesis
- Fail to reject the null hypothesis $H_0$: conclude that the data could possibly support the null hypothesis

We decide between these two outcomes using a *test statistic*, which is a function of the observed data.

The value of the test statistic determines the decision.

# Review of Hypothesis Testing Ideas/Vocabulary

There are two different errors we could make as a result of the hypothesis testing decision:

- Type I error: Reject $H_0$ when $H_0$ is in fact true
- Type II error: Fail to reject $H_0$ when $H_0$ is not true.

| | **Truth** | |
|---|---|---|
| **Decision** | $H_0$ | NOT $H_0$ |
| Reject $H_0$ | Type I Error | Correct Decision |
| Fail to Reject $H_0$ | Correct Decision | Type II Error |

# Review of Hypothesis Testing Ideas/Vocabulary

Hypothesis tests are designed to control the probability of making a Type I error:

$$
\begin{aligned}
\alpha &= P(\text{Type I error}) \\
&= P(\text{Reject } H_0 \text{ when } H_0 \text{ is true}) \\
&= P_{H_0}(\text{Reject } H_0)
\end{aligned}
$$

This probability $\alpha$ is called the *significance level* of the test.

# Review of Hypothesis Testing Ideas/Vocabulary

The probability of making a Type II error depends on the <u>true value</u> of the parameter being tested.

Often instead of talking about the probability of a Type II error, we talk about the *power* of a test procedure at a specific value $\mu_A$ of the parameter tested

$$\begin{aligned}\text{Power}(\mu_A) &= P(\text{Reject } H_0 \text{ when } \mu_A \text{ is true}) \\ &= P_{\mu_A}(\text{Reject } H_0)\end{aligned}$$

When comparing tests with the same significance level, we prefer a test with ***higher power***, since that means the test will make the correct decision more often.

# Univariate Hypothesis Test for Population Mean

Recall the $t$-test:

- Setting: $X_1, \dots, X_n$ iid from a population with unknown mean $\mu$ (and unknown variance $\sigma^2$)
- Null hypothesis:
$$H_0: \mu = \mu_0$$
- Alternative hypothesis:
$$H_A: \mu \neq \mu_0$$
- Test statistic:
$$t(\mu_0) = \frac{\sqrt{n}(\bar{X} - \mu_0)}{s}$$
- Decision:
Reject $H_0$ at level $\alpha$ for $|t(\mu_0)| > t_{(n-1)}(\alpha/2)$

# Univariate Hypothesis Test for Population Mean

Recall the $t$-test:

- Setting: $X$
  mean $\mu$ (

Throughout this course, we will only be considering tests that ***do not assume that the population variance (or covariance matrix) is known*** (the textbook *does* cover the case where the variance is known, however).

This is because in practice, the variance/covariance is pretty much ***never*** known if the population mean is not known.

- Null hypo
- Alternati
- Test statistic:

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

- Decision:

Reject $H_0$ at level $\alpha$ for $|t(\mu_0)| > t_{(n-1)}(\alpha/2)$

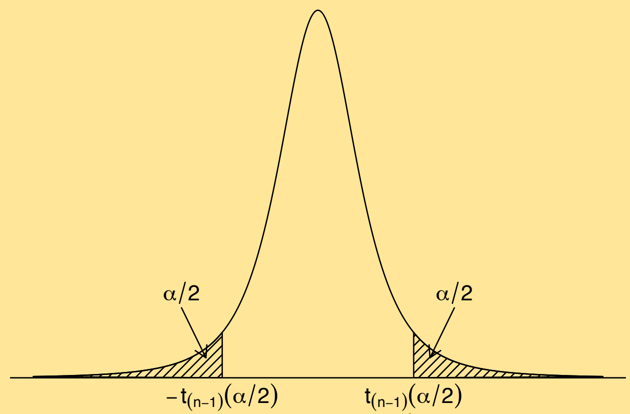
# Univariate Hypothesis Test for Population Mean

Recall the $t$-test

- Setting: $X_1$, ...
  mean $\mu$ (and
- Null hypothes
- Alternative hy
- Test statistic:

$t(\mu_0)$ ... $s$

The *critical value* $t_{(n-1)}(\alpha/2)$ is the upper $\alpha/2$ quantile of a $t$-distribution with (n-1) degrees of freedom:




- Decision:

$$\text{Reject } H_0 \text{ at level } \alpha \text{ for } |t(\mu_0)| > t_{(n-1)}(\alpha/2)$$

# Univariate Hypothesis Testing Example

**<u>Example:</u>** We have measured body temperature on $n$ = 12 healthy subjects.

| Subject | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Body Temp. | 94.3 | 95.2 | 98.2 | 96.8 | 99.3 | 98.5 | 99.4 | 97.6 | 99.2 | 98.3 | 97.9 | 98.0 |

We would like to use this data to perform a level $\alpha = 0.05$ test of the null hypothesis that the population mean (average) body temperature is 98.6° F.

$$H_0: \mu = 98.6$$

$$H_A: \mu \neq 98.6$$

# Univariate Hypothesis Testing Example

**Example:** We have measured body temperature on $n$ = 12 healthy subjects.

| Subject | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Body Temp. | 94.3 | 95.2 | 98.2 | 96.8 | 99.3 | 98.5 | 99.4 | 97.6 | 99.2 | 98.3 | 97.9 | 98.0 |

- Test statistic:

$$t(98.6) = \frac{\sqrt{12}(97.725 - 98.6)}{\sqrt{2.518}} = -1.910$$

- Critical value:

$t_{(11)}(0.025)$ = `qt(1-0.025, 11)` = 2.201

- Decision: Since $|t(98.6)| < t_{(11)}(0.025)$, we **fail to reject** the null hypothesis that the population mean is 98.6.

# Univariate Hypothesis Testing Example

**<u>Example:</u>** We have measured body temperature on $n$ = 12 healthy subjects.

| Subject | 1 | | 5 | 6 | 7 | 8 | | 12 |
|---|---|---|---|---|---|---|---|---|
| Body Temp. | 94.3 | | 99.3 | 98.5 | 99.4 | 97.6 | | 8.0 |

Sample mean
$\bar{X} = 97.735$

Sample variance
$s^2 = 2.518$

- Test statistic:

$$t(98.6) = \frac{\sqrt{12}(97.725 - 98.6)}{\sqrt{2.518}} = -1.910$$

- Critical value:

$t_{(11)}(0.025)$ = `qt(1-0.025, 11)` = 2.201

- Decision: Since $|t(98.6)| < t_{(11)}(0.025)$, we **fail to reject** the null hypothesis that the population mean is 98.6.

# Multivariate Hypothesis Testing

Now, with multivariate data $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$, we want to test hypotheses regarding a *mean vector*

$$\boldsymbol{\mu} = \left[\mu_1, \mu_2, \ldots, \mu_p\right]^T$$

- Null hypothesis:

$$H_0\colon \boldsymbol{\mu} = \boldsymbol{\mu}_0$$

- Alternative hypothesis:

$$H_A\colon \boldsymbol{\mu} \neq \boldsymbol{\mu}_0$$

How can this be done?

# Multivariate Hypothesis Testing Example

**Example:** We have measured body temperature and pulse on $n = 12$ healthy subjects.

| Subject | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Body Temp. | 94.3 | 95.2 | 98.2 | 96.8 | 99.3 | 98.5 | 99.4 | 97.6 | 99.2 | 98.3 | 97.9 | 98.0 |
| Pulse | 72 | 75 | 76 | 78 | 73 | 79 | 77 | 70 | 79 | 72 | 87 | 76 |

We would like to use this data to perform a level $\alpha = 0.05$ test of the null hypothesis that the population mean vector is (98.6° F, 77 bpm).

$$H_0: \boldsymbol{\mu} = (98.6, 77)$$

$$H_A: \boldsymbol{\mu} \neq (98.6, 77)$$

# Multivariate Hypothesis Testing

One idea for testing a hypothesis $H_0\colon \boldsymbol{\mu} = \boldsymbol{\mu}_0$:

- Just test each variable's mean separately: Using only the observations for variable *j*, perform a level $\alpha$ test of $H_{0j}\colon \mu_j = \mu_{0j}$, where
  - $\mu_j$ is the true (unknown) population mean of the *j*th variable
  - $\mu_{0j}$ is the *j*th element of the hypothesized mean vector $\boldsymbol{\mu}_0$
- Reject $H_0$ if ***any*** of the $H_{0j}$ are rejected.

Will this work? Will this approach provide a level $\alpha$ test of $H_0\colon \boldsymbol{\mu} = \boldsymbol{\mu}_0$?

# Multivariate Hypothesis Testing

Remember, our goal is to achieve

$$P_{H_0}(\text{Reject } H_0\text{: } \boldsymbol{\mu} = \boldsymbol{\mu}_0) = \alpha$$

If we test each individual variable at level $\alpha$, then we will ***inflate our Type I error probability***, because we are giving ourselves multiple opportunities to make a Type I error: this is a <u>*multiple testing*</u> scenario.

Therefore, this approach ***will not*** provide a level $\alpha$ test of $H_0\text{: } \boldsymbol{\mu} = \boldsymbol{\mu}_0$.

# Multivariate Hypothesis Testing

We could correct for the *multiple testing* issue using a *Bonferroni correction*:

- Test each individual variable's mean at level
$$\alpha^* = \frac{\alpha}{p}$$
(where $p$ is the number of variables measured, as usual).

Then we would have:

$$P_{H_0}(\text{Reject } H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0) = P_{H_0}\left(\text{Reject at least one of } H_{0j}\right)$$
$$\leq \sum_{j=1}^{p} P_{H_{0j}}\left(\text{Reject } H_{0j}\right) = \sum_{j=1}^{p} \frac{\alpha}{p} = \alpha$$

So the probability of a Type I error would be controlled: guaranteed to be less than or equal to the desired level $\alpha$.

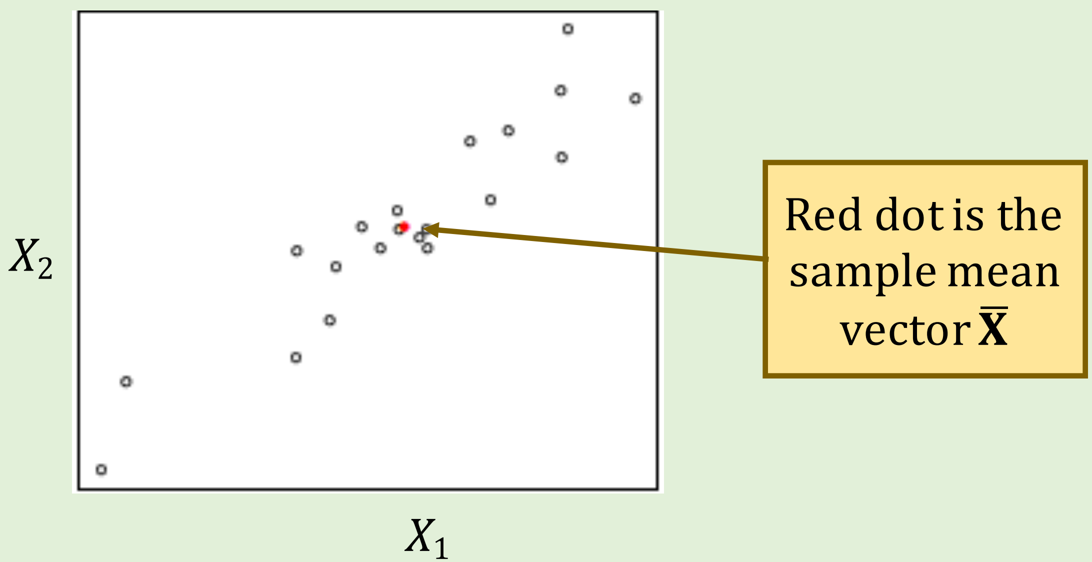
# Multivariate Hypothesis Testing

However, in many cases performing individual tests using a Bonferroni correction and combining the results will not be a desirable solution:

- This procedure is too *conservative* for most settings: it rejects a true null hypothesis with probability strictly less than the desired significance level $\alpha$.
- Conservatism generally implies a lack of power: this test will not reject the null in many settings where the null is in fact false.

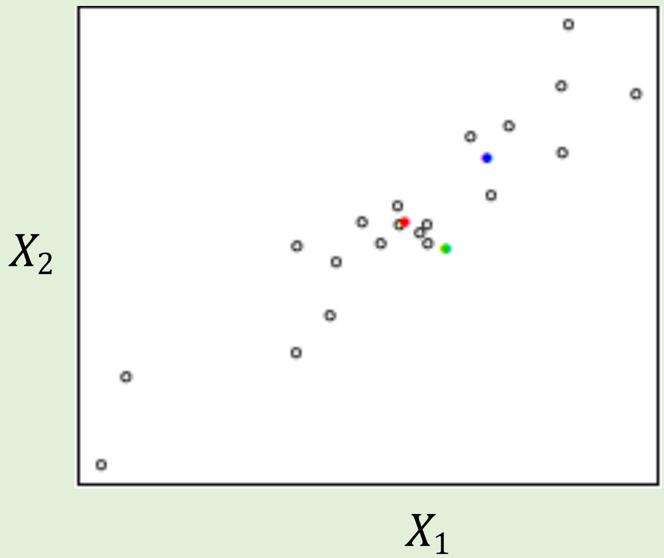
# Multivariate Hypothesis Testing: Example

Suppose we have the bivariate data represented in this scatterplot:



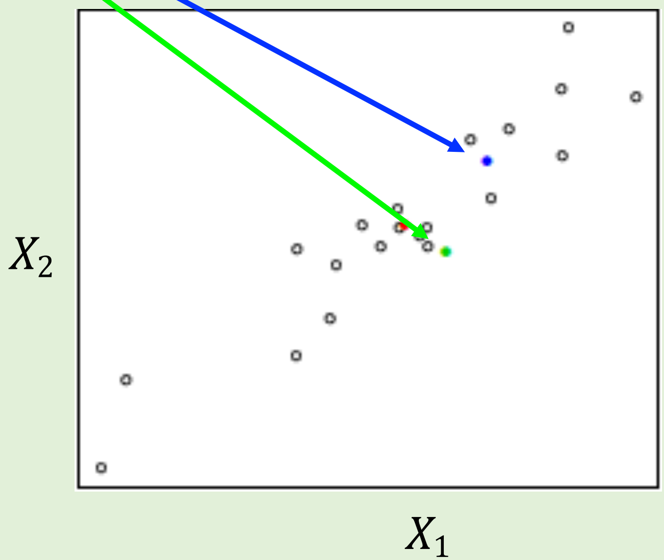
# Multivariate Hypothesis Testing: Example

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



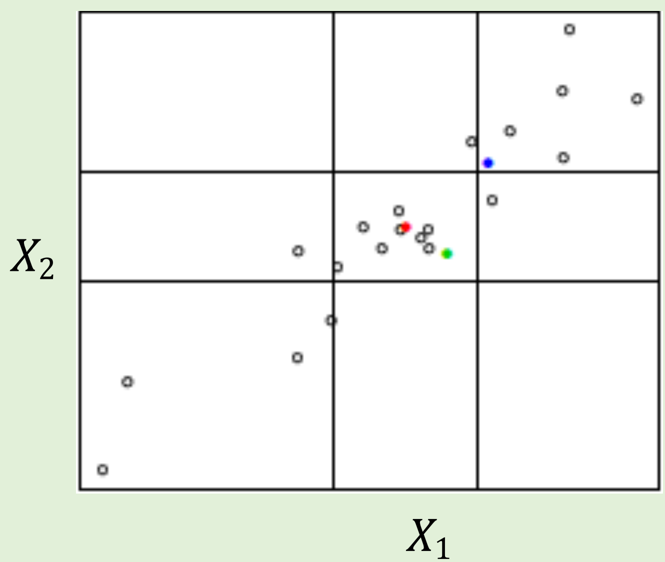
# Multivariate Hypothesis Testing: Example

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



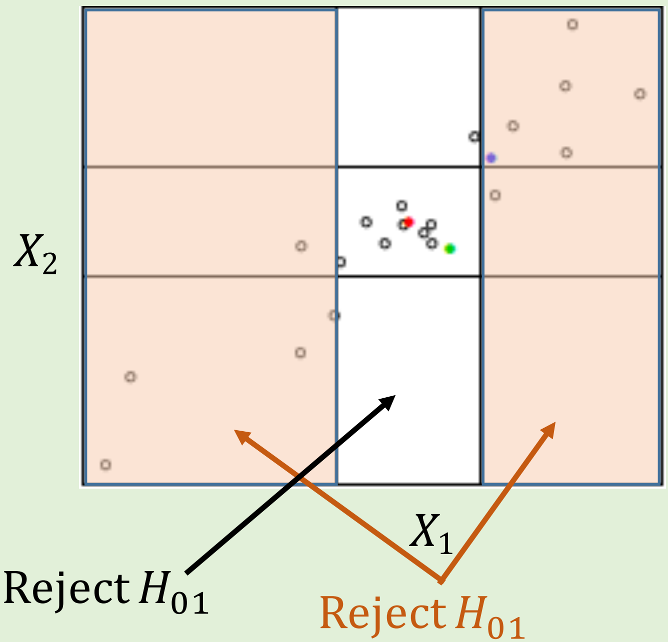
# Multivariate Hypothesis Testing: Example

If we performed univariate tests of each variable's mean separately, the black lines indicate the range of population mean values that would NOT be rejected:



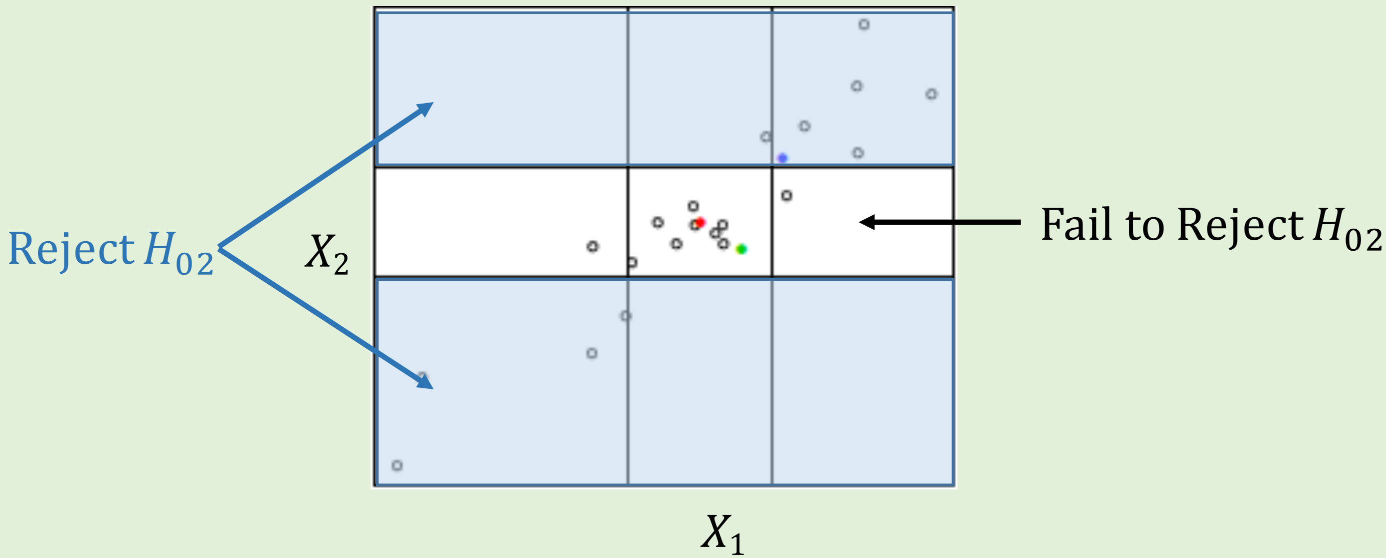
# Multivariate Hypothesis Testing: Example

If we performed univariate tests of each variable's mean separately, the black lines indicate the range of population mean values that would NOT be rejected:



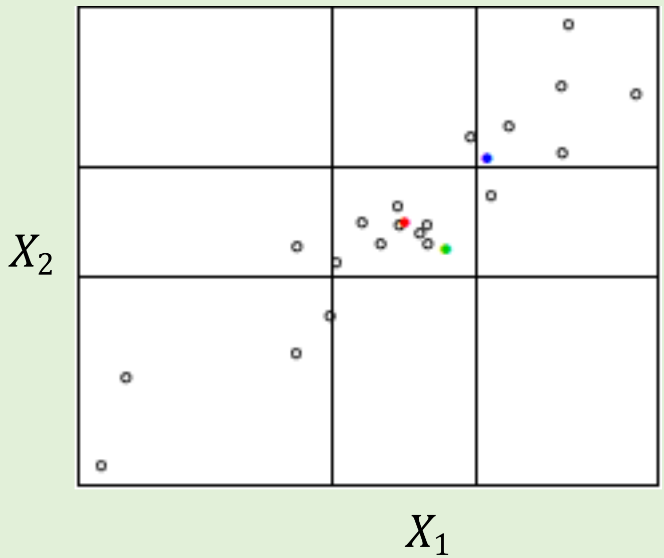
# Multivariate Hypothesis Testing: Example

If we performed univariate tests of each variable's mean separately, the black lines indicate the range of population mean values that would NOT be rejected:



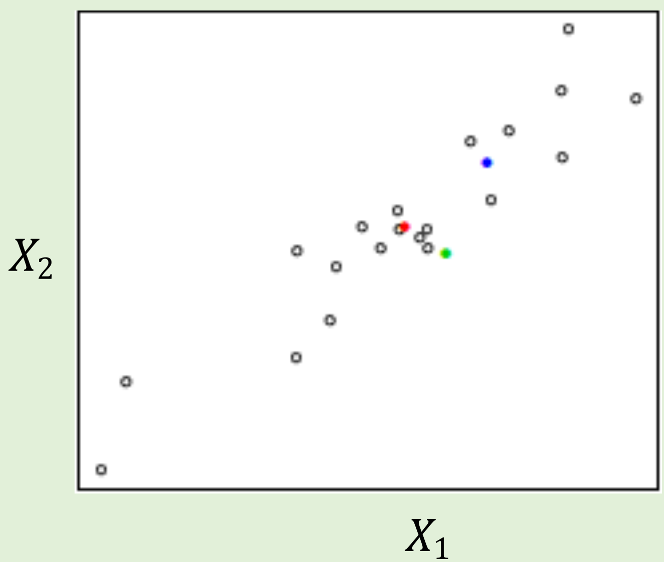
# Multivariate Hypothesis Testing: Example

Therefore, if we test each variable's mean separately, we would *reject* the blue point, but *fail to reject* the green point as plausible population mean vectors.



# Multivariate Hypothesis Testing: Example

However, we might convincingly argue that actually the blue point looks more reasonable than the green point: it is more centered within the data cloud.



# Multivariate Hypothesis Testing: Example

The separate univariate tests do not take into account the relationship *between* the variables, but <u>that relationship can give us useful information</u>.