MONASH University

Information Technology

FIT1006 Business Information Analysis

Lecture 18 Hypothesis Testing (Part 2)

Topics covered:

- p-value
- Type I and II errors
- The significance and power of a test.

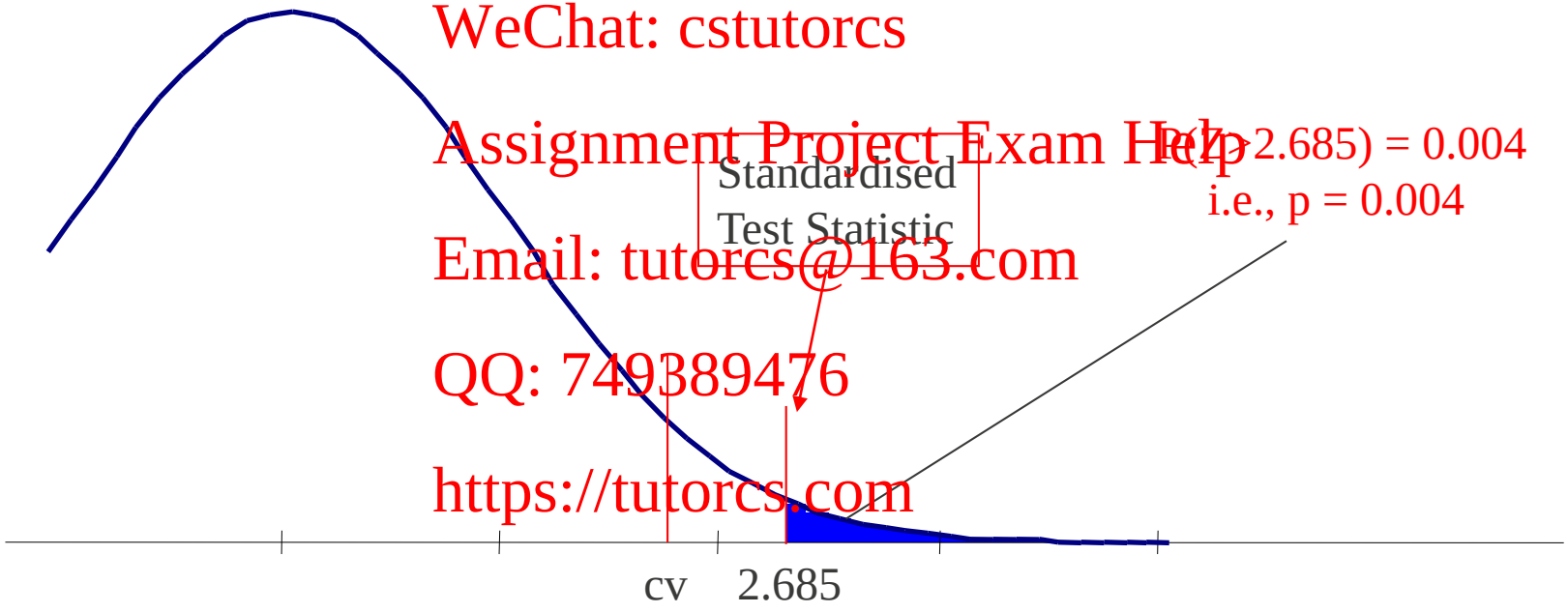
The Steps of Hypothesis Testing

1. Decide on a null hypothesis H_0.
2. Decide on an alternative hypothesis H_1.
3. Decide on a significance level.
4. Calculate the appropriate test statistic.
5. Find from tables the corresponding tabulated test statistic.
6. Compare calculated and tabulated test statistics and decide whether to accept or reject the null hypothesis.
7. State the conclusion and assumptions of the test.

(Source: Rees, D.G. Essential Statistics, Chapman and Hall 1995.)

p-value

- The p-value of [illegible] s the smallest value of the critical region t[illegible]s to the rejection of H_0. For example, if a te[illegible]he following test statistic:

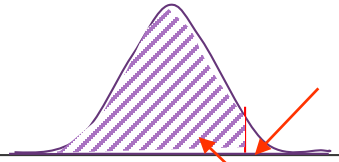


Calculating the p-value

- Modified Questio st lecture:
- A hypothesis test opulation mean when the population variance is known. (Note: unrealistic situation)
- The Axle manufacturing company has been making axles for a long time and kept records for every axle produced.
- Population parameters are μ = 90mm and σ = 2.5mm.
- A sample of 100 axles from new machine has mean = 90.5.
- Is the new machine making parts with same average length required (90mm) – or are they longer than this?
- What is the p-value of the test?

Question 1

From the tables 2) =



$P(Z>2) = 1 - 0.9772$
$= 0.0228$

This area = 0.9772; z = 2.0
(one tail)

A. 0.0114

✓ B. 0.0228

C. 0.4886

D. 0.9772

Cumulative Probabilities for the Standard Normal Distribution

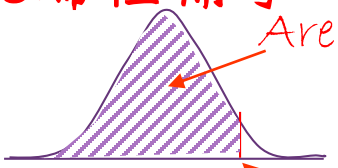
Table gives P(Z<z) for Z = N(0,1)

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952

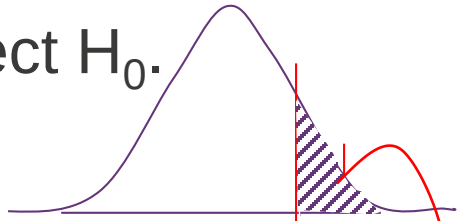
Given in Slide 5:
$\mu = 90mm$ and $\sigma = 2.5mm$

Solution

1. H_0, $\mu = 90mm$
2. H_1, $\mu > 90mm$ (1-sided experiment)
3. Significance = 0.01.
4. The test statistic, $\bar{x} = 90.5$. We calculate Z_x. *(standardising)* $Z = (90.5 - 90)/(2.5/\sqrt{100}) = 2$.
5. From tables the calculated critical value is 2.33.
6. We see that $2 < 2.33$ and thus do not reject H_0.
7. Conclude …
8. *p-value* of the test is 0.0228 – this is the highest level of significance at which H_0 will be rejected.



One tail critical value for $\alpha = 0.01$, from table: $z = 2.33$



$P(Z > 2) = 1 - 0.9772$
$= 0.0228$

… visually…

Interpreting the p-value:

- A small p-value indicates tha ample evidence to support th alternative hypothesis.
- A big p-value indicates that there is little evidence to support the alternative hypothesis.

Since, 2 < 2.33
→ do not reject H_0

Test statistic = 90.5
Standardised test statistic = 2

$P(Z > 2) = 1 - 0.9772$
$= 0.0228$
→ p-value = 0.0228

$\mu = 90$

For $\alpha = 0.01$, one tail;
critical value = 2.33

Errors in Hypothesis Testing

- We can make two [illegible] when testing hypotheses
  - A Type I error [illegible] hen the null hypothesis is correct, but is rejected.
  - A Type II error occurs when the null hypothesis is incorrect but is not rejected.

	H_0 is Not Rejected	H_0 is Rejected
H_0 is Correct	Correct Decision (1 - α)	Type I Error α
H_1 is Correct	Type II Error β	Correct Decision (1 - β)

…

- The distinction between type I and II errors is easier to understand using criminal law analogy.

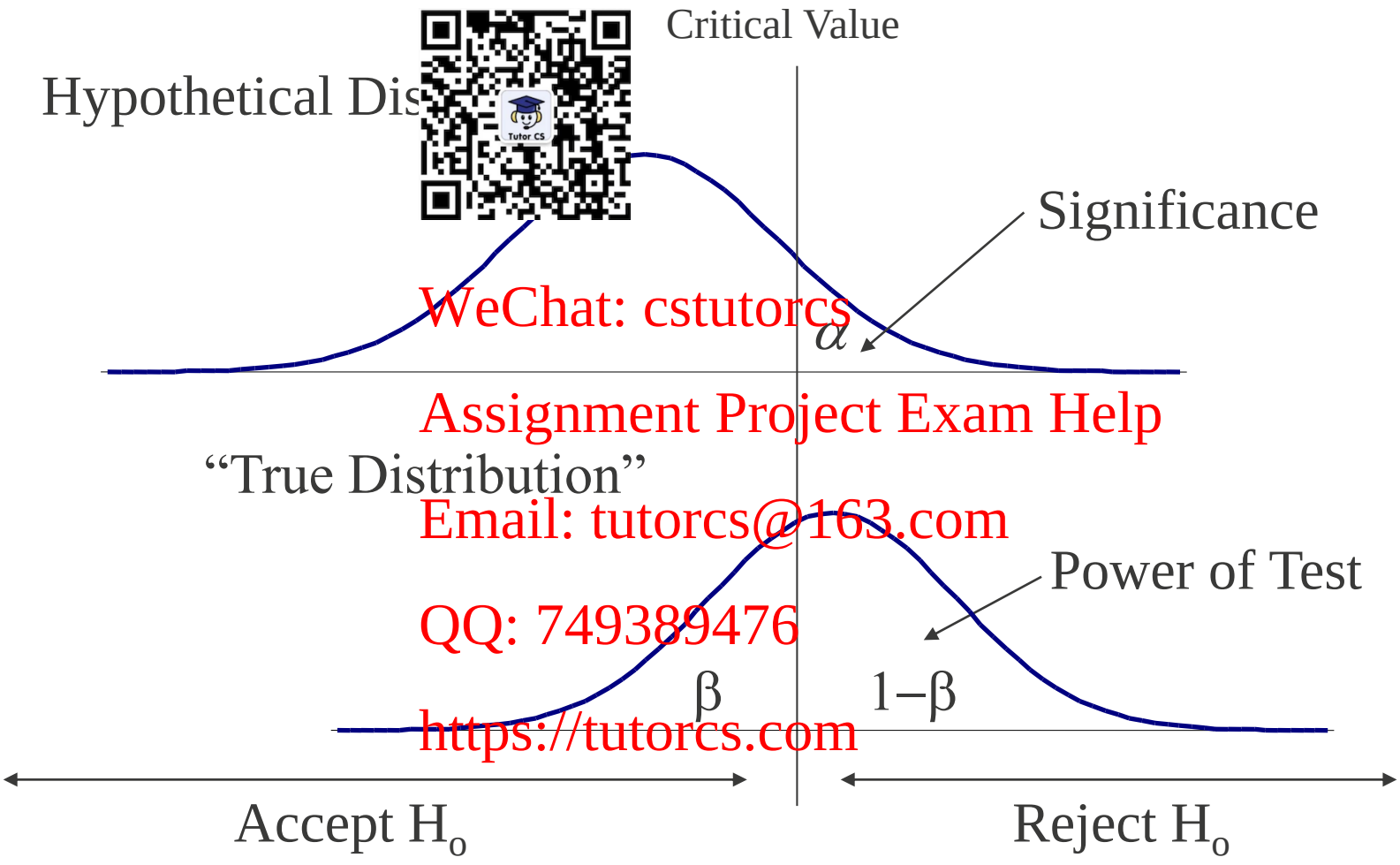
	Person Acquitted	Person Convicted
Person is Innocent	Correct Decision	Type I Error
Person is Guilty	Type II Error	Correct Decision

- When we attempt to minimize the type I error, the consequence is an increase in the type II error.

The Power of the Test

- The probability of making a Type II error is β. If the null hypothesis is incorrect, the probability of avoiding a Type II error is (1-β), this is the power of the test.
- α and β are interrelated. Increasing α reduces β and *vice versa*. To reduce both, sample size needs to be increased.
- Only α or β (usually α) can be chosen. The true value of β also depends on the actual value of the population parameter – which is often unknown.

The Power of the Test ctd..



Calculating β

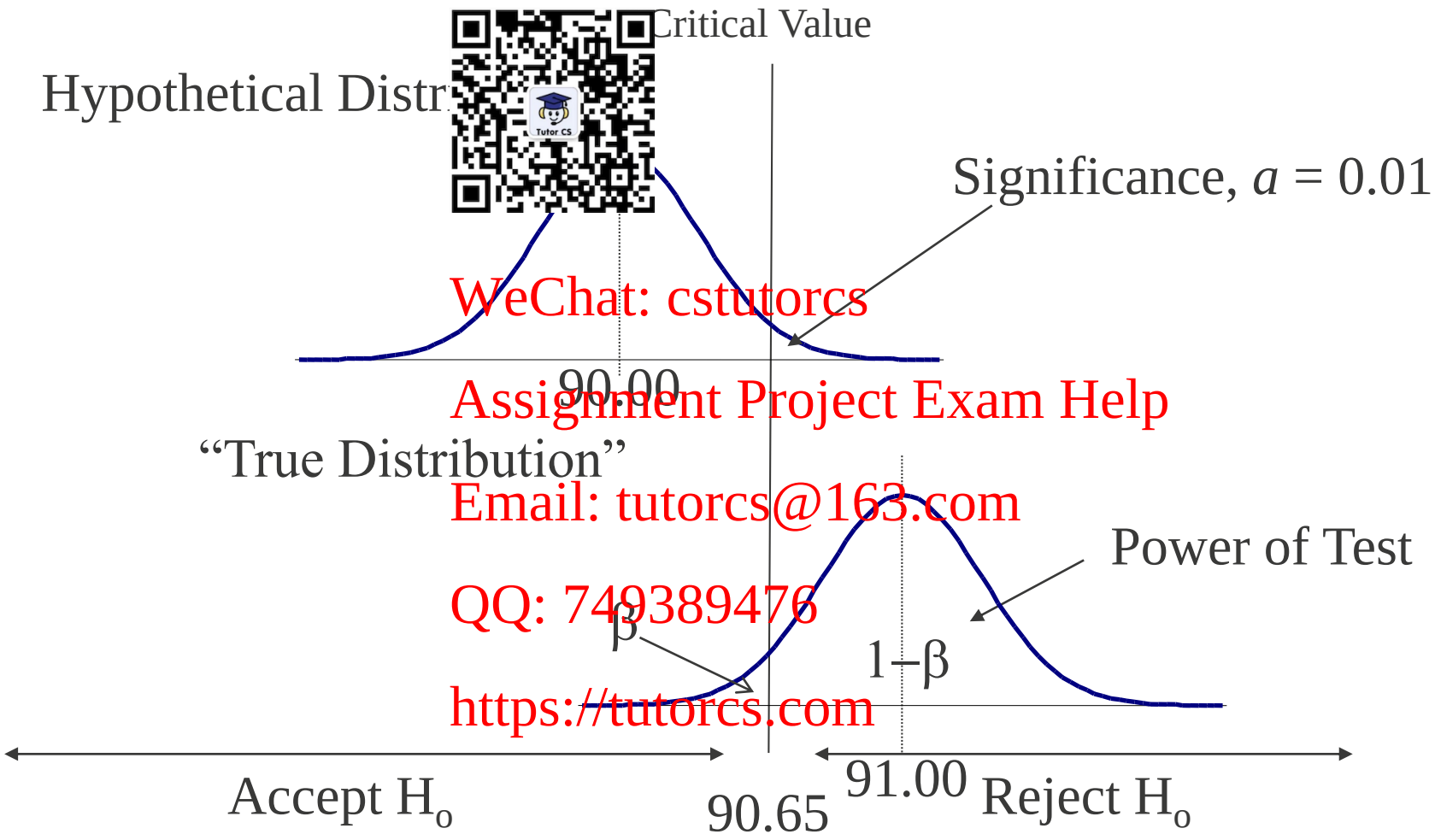
- Using the same e

  - The Axle mar g company has been making axles for a long time t records for every axle produced.
  - Population parameters are μ = 90mm and σ = 2.5mm.
  - A sample of 100 axles from new machine has mean = 92.7.
  - Is the new machine making parts with same average length required (90mm)?
  - Assume a 1% significance.
  - What is the power of the test if it eventuates that the new machines produces axles with a mean length of 91mm?

Solution

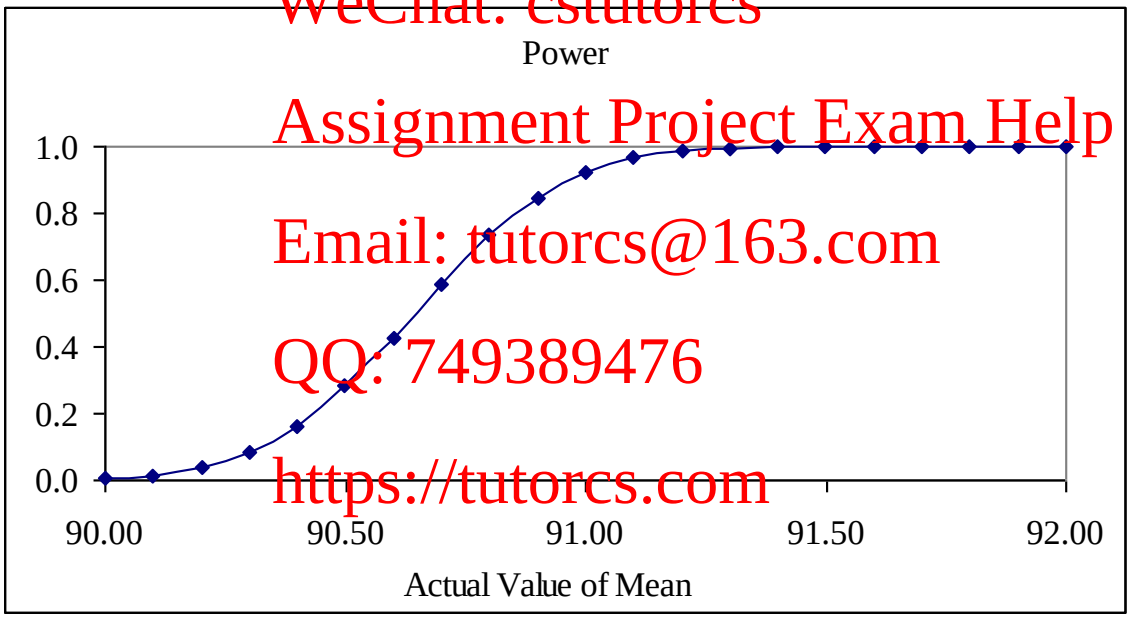
- From our previous calculations
- H_0, $\mu = 90$mm
- H_1, $\mu \neq 90$mm (a two tailed experiment)
- Significance = 0.01.
- From tables the critical values are $\pm 2.58\sigma_{\bar{x}}$
- Use upper critical value since alternative is > H_0
- Calculate upper critical value = 90 + 2.58(0.25) = 90.65
- P(x < 90.65) when $X \sim N(91, 0.25^2)$ is P(z < -1.42) = 0.08
- This is the probability that we fail to reject the null hypothesis, when the population mean for the new axle is 91mm. The power of the test = 0.92

The Power of the Test ctd..



Power Curves

- To see how the power of a test changes as the value of the alternative hypothesis changes, we can calculate the power for a range of values and view the resulting curve.



Example

- From last lectu
- It is claimed th urne families are spending more than $150 per week on food and grocery items on average. A sample of 15 families was conducted and the amount spent each week was recorded. Do these results support this claim? (assume a 1% significance)
- Weekly food and grocery expenditure ($)
- 156, 234, 199, 78, 256, 189, 221, 49, 220, 178, 120, 290, 97, 177, 231.
- What is the power of the test if the average cost of groceries is $160?

Summary Statistics from Excel

Mean	179.7
Standard Error	17.7
Median	189.0
Mode	#N/A
Standard Deviation	68.5
Sample Variance	4697.0
Kurtosis	-0.5
Skewness	-0.5
Range	241.0
Minimum	49.0
Maximum	290.0
Sum	2695.0
Count	15.0

Question 2

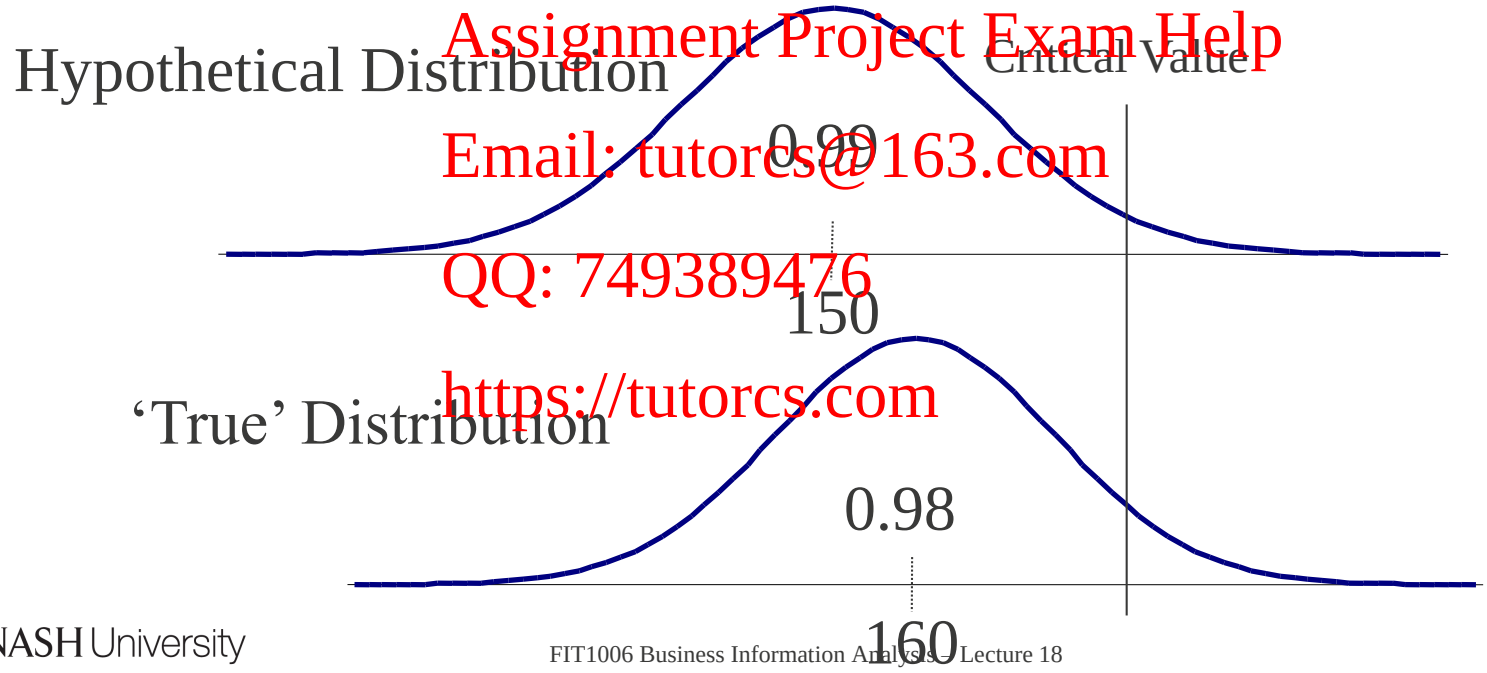
The power of th s:

A. 0.99

B. 0.01

C. 0.98

✓ D. 0.02



Question 3

For a one-sided [illegible] 15 observations the *t* Statistic for a 1% significance is:

A. 2.6025

B. 2.9467

✓ C. 2.6245

D. 2.9768

Critical Values of the t Distribution

Table gives upper critical values only

				a				
n	**0.300**	**0.200**	**0.150**	**0.100**	**0.050**	**0.025**	**0.010**	**0.005**
12	0.5386	0.8726	1.0832	1.3562	1.7823	2.1788	2.6810	3.0545
13	0.5375	0.8702	1.0795	1.3502	1.7709	2.1604	2.6503	3.0123
14	0.5366	0.8681	1.0763	1.3450	1.7613	2.1448	2.6245	2.9768
15	0.5357	0.8662	1.0735	1.3406	1.7531	2.1314	2.6025	2.9467
16	0.5350	0.8647	1.0711	1.3368	1.7459	2.1199	2.5835	2.9208
17	0.5344	0.8633	1.0690	1.3334	1.7396	2.1098	2.5669	2.8982
18	0.5338	0.8620	1.0672	1.3304	1.7341	2.1009	2.5524	2.8784

...

- H_0, μ = \$150
- H_1, μ > \$150 (a c experiment)
- Significance = 0.
- From tables the critical value $T_{(0.01)(\nu=14)}$ is $2.624\sigma_x$
- Calculate upper critical value: 150 + 2.624(17.7) = 196.5
- Assume alternative distribution Normal, new μ, original SE.
- P(x < 196.5) when $X \sim N(160, 17.7^2)$ = P(z < 2.061) = 0.98
- This is β, the probability that we fail to reject the null hypothesis, when the population mean for expenditure is \$160.
- The power of the test is 0.02.

Other Hypothesis Tests

- We have looked at hypothesis tests for the population mean and a population proportion.
- There are a range of other hypothesis tests which are regularly used, these include tests for:
  - The difference of means
  - Variance
  - Distribution Shape
  - Independence.

We're not going to cover this.

Reading/Questions (Selvanathan)

- Reading: Hypothe ng
  - 7th Ed. Se 3, 12.5, 12.6.

- Questions: Hypothesis Testing

  - 7th Ed. Questions 12.1, 12.19, 12.25, 12.26, 12.56, 12.59, 12.65, 12.66, 12.67, 12.70, 12.72, 12.74.