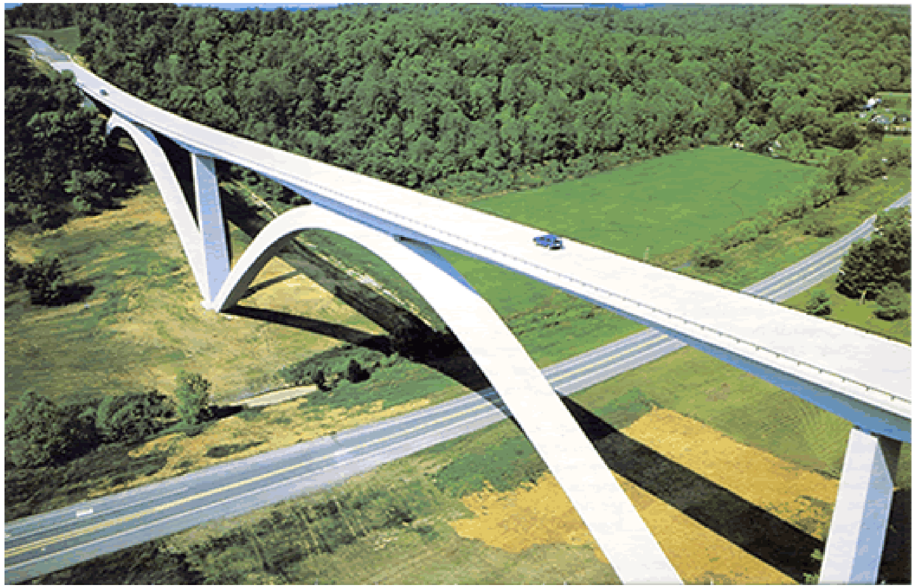
Quantitative Risk Analysis Using Hypothesis Testing

COORDINATOR:
Dr Lihai Zhang
Infrastructure Engineering Department (Room B307)
lihzhang@unimelb.edu.au



Quantitative Analysis – Quality Risks

Consider a machine that makes steel bars for use in building construction. The specification for the diameter of the bars is 2.0+/-0.1 cm. During the last hour, the machine has made 1000 rods.

The quality engineer draws a random sample of 50 rods, measures them, and finds that 46 of them (92%) meet the diameter specification.

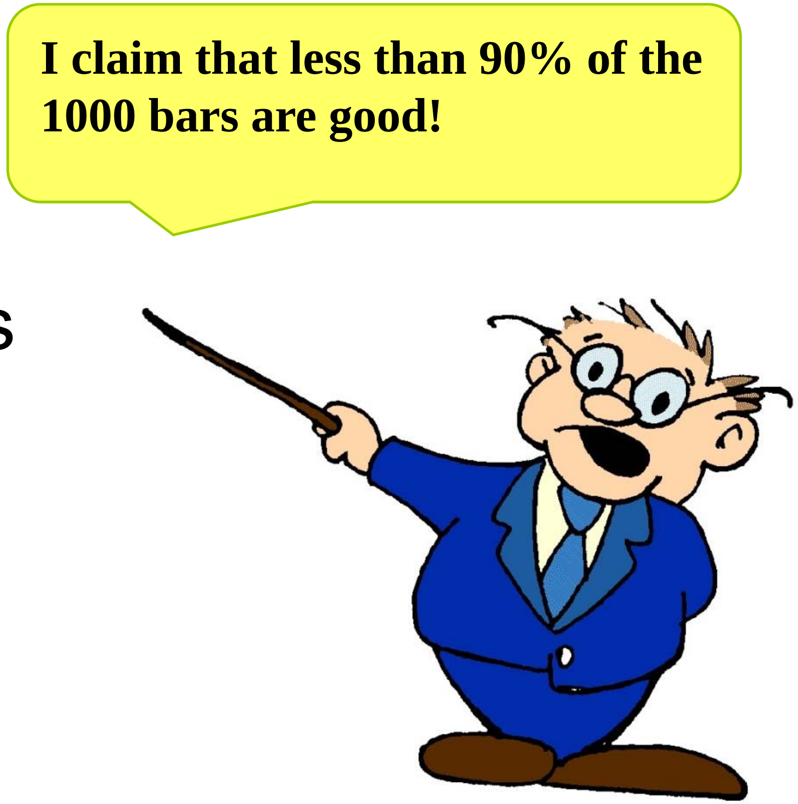
It is unlikely that the sample of 50 bars represents the population of 1000 perfectly !

Questions

- The engineer wants to be fairly certain that the good steel bars >= 90%; otherwise the machine will be shut down for recalibration. How certain can he be that at least 90% of the 1000 bars are good?
  - Requires a Hypothesis Test

What is a Hypothesis?

- A hypothesis is a claim (assumption) about the population parameter
  – Examples of parameters are population mean or proportion
  – The parameter must be identified before analysis



The Null Hypothesis, H_0

- States the assumption (numerical) to be tested
  - e.g. The average number of bushfires in Melbourne every year is more than three ($H_0: X \leq 3$)
- Always about a population parameter, not about a sample statistic

The Null Hypothesis, H_0

- Begins with the assumption that the null hypothesis is true
  - Similar to the notion of innocent until proven guilty
- Refers to the status quo
- Always contains the **"="** sign
- May or may not be rejected

The Alternative Hypothesis, H_1

- Is the opposite of the null hypothesis
- Challenges the status quo
- Never contains the "=" sign
- May or may not be accepted

The average number of bushfires in Melbourne every year

Null Hypothesis	$H_0 : X \leq 3$
Alternative Hypothesis	$H_1 : X > 3$

Steps in Performing a Hypothesis Test

1. Define H_0
2. Assume H_0 to be true
3. Assess the strength of the evidence against H_0 using ***P*-value**
   - The ***P*-value** is the probability assuming H_0 to be true.
   - A rule of thumb: Reject H_0 whenever

$$P \leq \alpha = 0.05$$

Significant Level

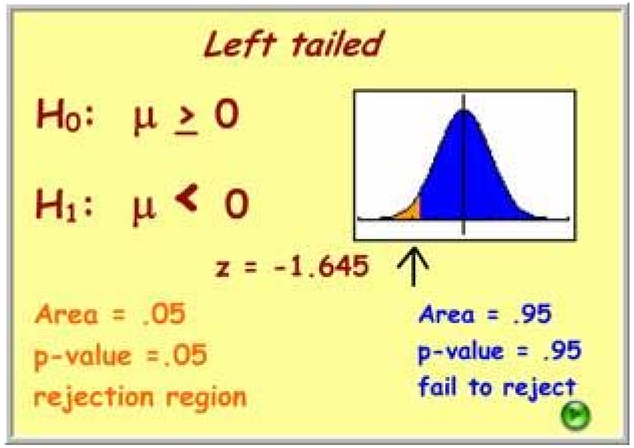
While this rule is convenient, it has no scientific basis!

Calculation of *P*-value

- For a large (e.g. $n > 30$) sample from a population, the ***P*-value** is an area under the normal curve.
- For a small (e.g. $n \leq 30$) sample from a population, the ***P*-value** is an area under the Student's *t* curve with n-1 degrees of freedom.
- The smaller the ***P*-value**, the stronger the evidence is against H_0.
- The larger the ***P*-value**, the more plausible H_0 becomes.

Large-Sample Tests for a Population mean

Compute the z-score: $$z = \frac{\overline{X} - \mu_0}{S / \sqrt{n}}$$



Compute the ***P*-value** :

Null Hypothesis	Alternative Hypothesis	*P* - value
$H_0 : \mu \le \mu_0$	$H_1 : \mu > \mu_0$	Area to the right of *z*
$H_0 : \mu \ge \mu_0$	$H_1 : \mu < \mu_0$	Area to the left of *z*
$H_0 : \mu = \mu_0$	$H_1 : \mu \ne \mu_0$	Sum of the areas in the tails cut off by z and -z

- If the null hypothesis is rejected, then we accept the alternative hypothesis.
- If the null hypothesis is not rejected, then we do not accept the alternative hypothesis.

Large-Sample Tests for a Population mean

Null Hypothesis	Alternative Hypothesis	*P* - value
$H_0 : \mu \le \mu_0$	$H_1 : \mu > \mu_0$	Area to the right of *z*

Example: $$H_0 : \mu \le 100$$

$$SE = \frac{S}{\sqrt{n}} = \frac{20}{\sqrt{40}} = 3.16$$

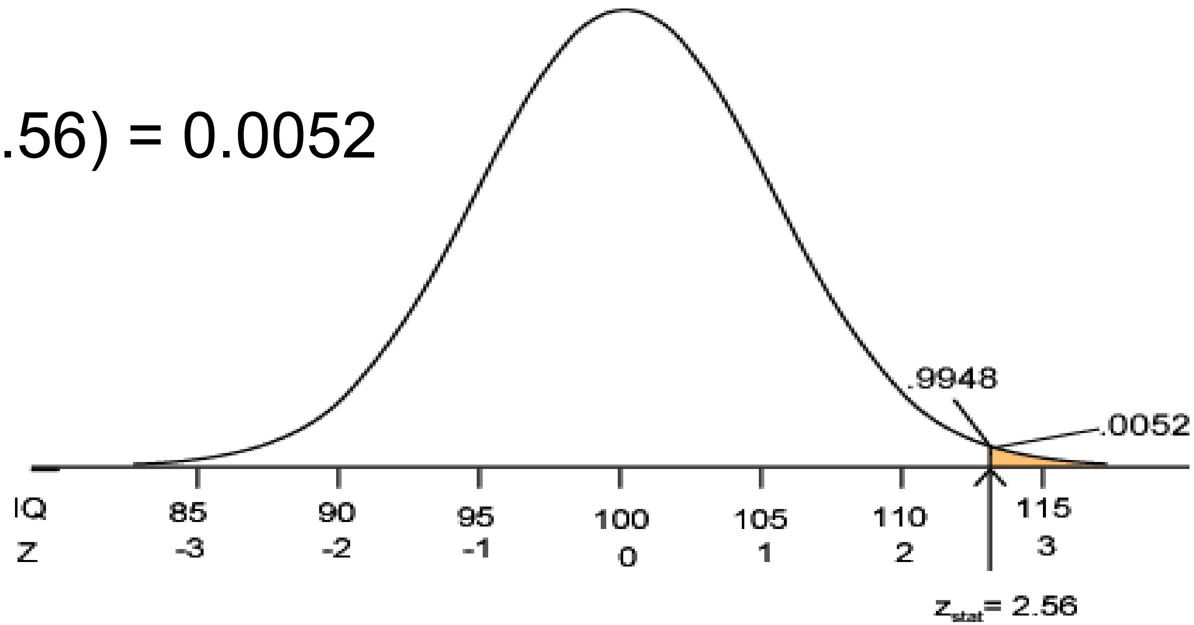
$$z = \frac{\bar{x} - \mu_0}{SE} = \frac{108.1 - 100}{3.16} = 2.56$$

Large-Sample Tests for a Population mean

Null Hypothesis	Alternative Hypothesis	*P* - value
$H_0 : \mu \leq \mu_0$	$H_1 : \mu > \mu_0$	Area to the right of *z*

***P*-value:**

$P = \Pr(Z \geq 2.56) = 0.0052$



$P = 0.0052 < 5\% \Rightarrow$ Strong evidence against H_0

Large-Sample Tests for a Population mean

Null Hypothesis	Alternative Hypothesis	*P* - value
$H_0 : \mu = \mu_0$	$H_1 : \mu \neq \mu_0$	Sum of the areas in the tails cut off by z and -z

Example:

$$H_0 : \mu = 100$$

$$SE = \frac{S}{\sqrt{n}} = \frac{20}{\sqrt{40}} = 3.16$$

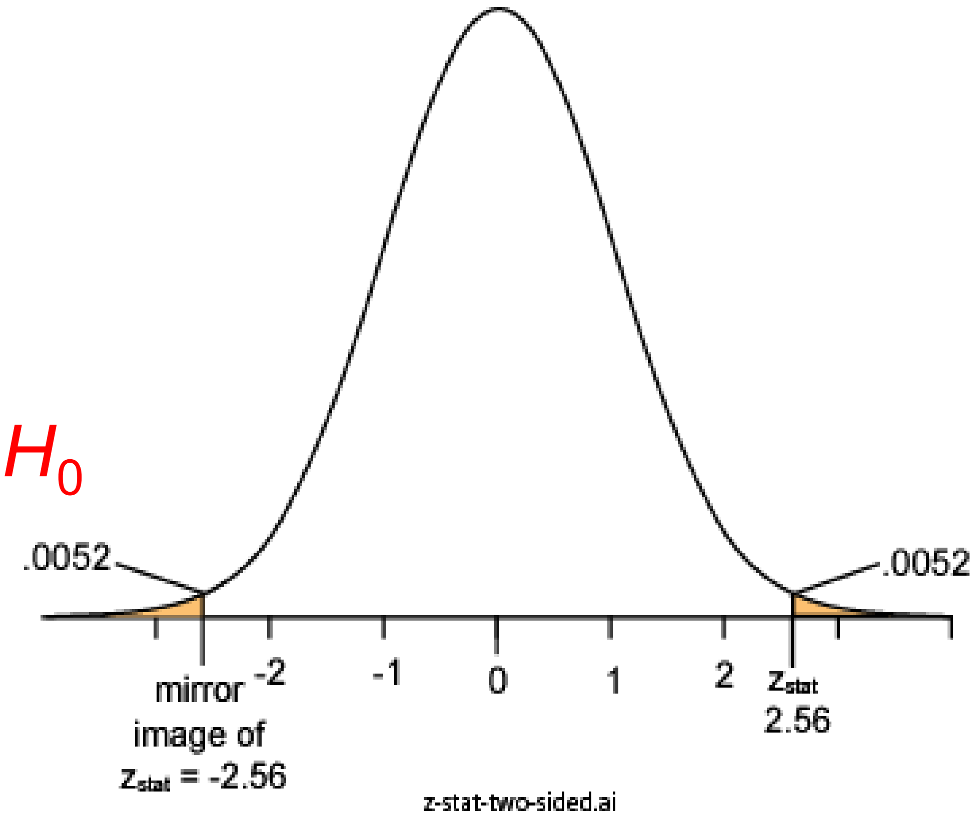
$$z = \frac{\bar{x} - \mu_0}{SE} = \frac{108.1 - 100}{3.16} = 2.56$$

Large-Sample Tests for a Population mean

Null Hypothesis	Alternative Hypothesis	*P* - value
$H_0 : \mu = \mu_0$	$H_1 : \mu \neq \mu_0$	Sum of the areas in the tails cut off by z and -z

- Two - sided *P* = 2 × 0.0052 = 0.0104 < 5%

Strong evidence against H_0



Example 1 (Environmental Risks)

Regulations require that the chlorine level in wastewater discharges be less than 100 μg/L. In a sample of 85 wastewater specimens, the mean chlorine concentration was 98 μg/L and the standard deviation was 20 μg/L. Let μ represent the mean chlorine level. A test is made of H_0: $\mu \geq 100$.

(a) Find the *P*-value.

(b) Do you believe it is plausible that the mean chlorine concentration is greater than or equal to 100 μg/L, or are you convinced that it is less?

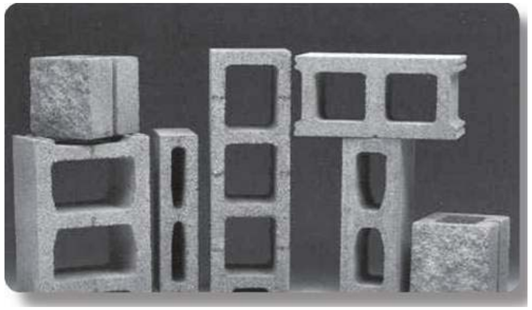
Solution:

Example 2 (Quality Risks)

A new concrete mix is being designed to provide adequate compressive strength for concrete blocks. The specification for a particular application calls for the blocks to have a mean compressive strength μ greater than 1350 kPa. A sample of 1000 blocks is produced and tested. Their mean compressive strength is 1356 kPa and their standard deviation is 70 kPa. A test is made of H_0: $\mu \leq 1350$ kPa.

(a) Find the *P*-value.

(b) Do you believe it is plausible that the blocks do not meet the specification, or are you convinced that they do?



Solution:

Example 3 (Quality Risks)

An inspector measured the fill volume of a simple random sample of 100 cans of juice that were labeled as containing 12 oz. The sample had mean volume 11.98 oz and standard deviation 0.19 oz. Let μ represent the mean fill volume for all cans of juice recently filled by this machine. The inspector will test H_0: $\mu = 12$.

(a) Find the *P*-value.

(b) Do you believe it is plausible that the mean filled volume is 12 oz?

Solution:

Small-Sample Tests for a Population mean

Compute the test statistic t:

$$t = \frac{\bar{X} - \mu_0}{S / \sqrt{n}}$$

Compute the ***P*-value** :

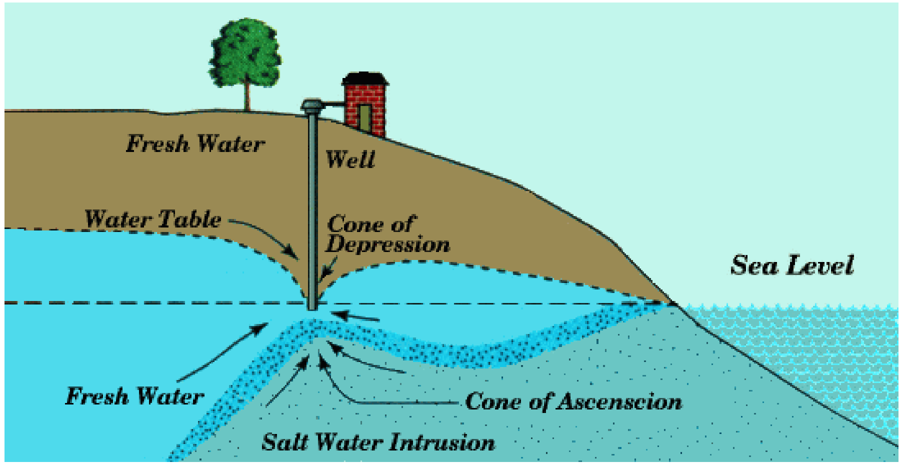
Null Hypothesis	Alternative Hypothesis	*P* - value
$H_0 : \mu \le \mu_0$	$H_1 : \mu > \mu_0$	Area to the right of *t*
$H_0 : \mu \ge \mu_0$	$H_1 : \mu < \mu_0$	Area to the left of *t*
$H_0 : \mu = \mu_0$	$H_1 : \mu \ne \mu_0$	Sum of the areas in the tails cut off by *t* and -*t*

Example 4 (Environmental Risks)

Measurement of groundwater concentrations of silica (SiO_2), in mg/L, were made at a sample of 12 wells in a certain city. The sample mean concentration was 61.3 and the standard deviation was 5.2.

(a) Can you conclude that the mean concentration of silica is greater than 60 mg/L?

(b) Can you conclude that the mean concentration of silica is less than 65 mg/L?



Solution:

Example 5 (Quality Risks)

As part of the quality-control program for a catalyst manufacturing line, the raw materials (alumina and a binder) are tested for purity. The process requires that the purity of the alumina be greater than 85%. A random sample from a recent shipment of alumina yielded the following results (in %):

93.2	**87.0**	**92.1**	**90.1**	**87.3**	**93.6**

A hypothesis test will be done to determine whether or not to accept the shipment.

(a) Compute the *P*-value?

(b) Should the shipment be accepted?

Solution:

Quality Risks

Can you conclude that the mean concrete strength (kPa) after six days is greater than the mean strength after three days?

	1	2	3	4	5	Average	STD
After 3 days	1341	1316	1352			**1336**	**18**
After 6 days	1376	1373	1366	1384	1358	**1371**	**9**

Dependent and Independent Samples

Are the samples from two groups are **Dependent** or **Independent** ?

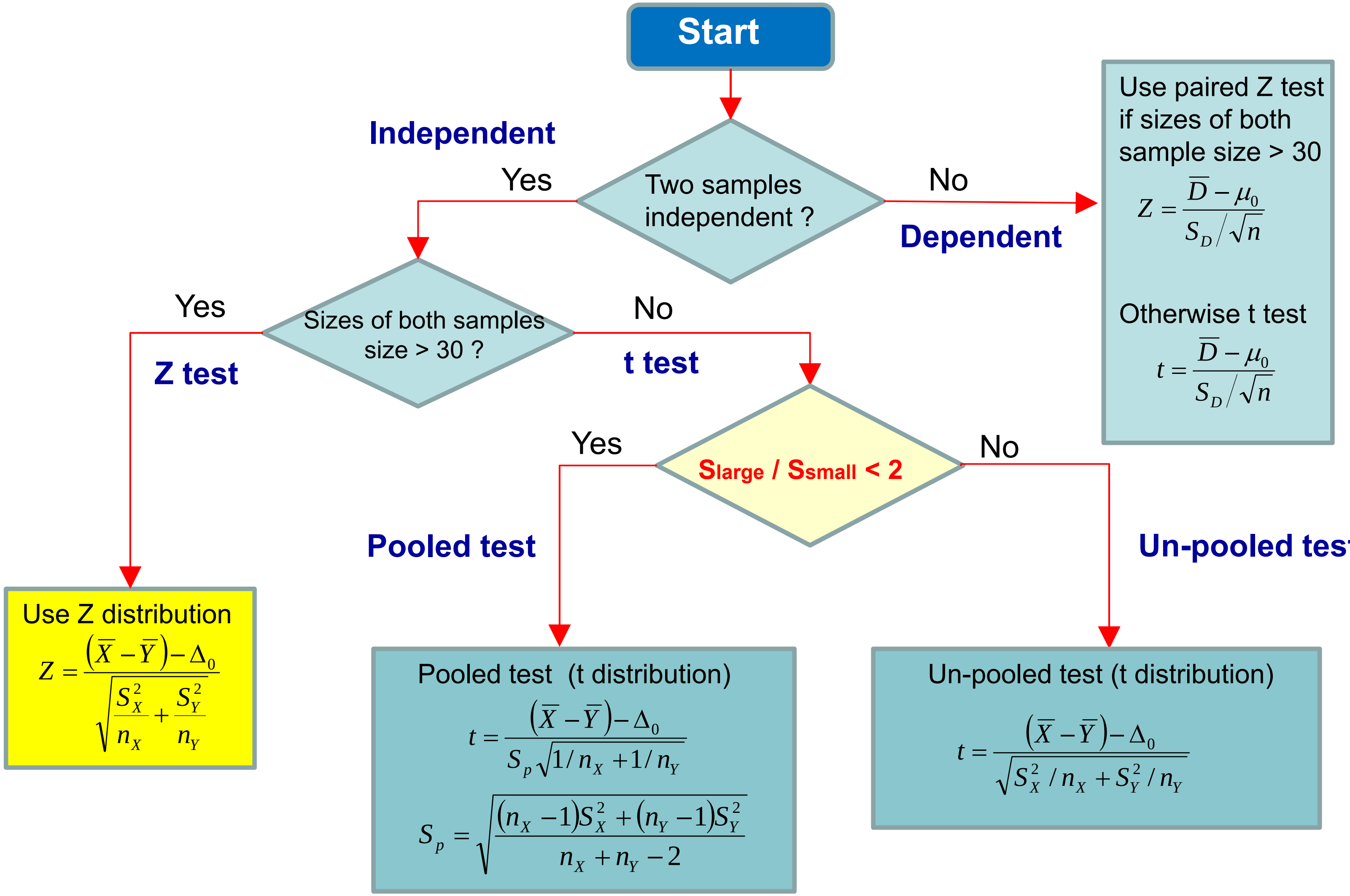
Dependent samples

- The difference between the number of floods in Melbourne for the past 3 years and 5 years (Samples are from the same population, i.e. Melbourne).

Independent samples

- The difference between the number of floods in Sydney and Melbourne for the past 3 years and 5 years (Samples are from two unconnected populations, i.e. Sydney and Melbourne)

Hypothesis Testing for the Difference Between Two Means



Independent samples: Large-Sample Tests for the Difference Between Two Means

Compute the z-score: $$Z = \frac{\left(\overline{X} - \overline{Y}\right) - \Delta_0}{\sqrt{\frac{S_X^2}{n_X} + \frac{S_Y^2}{n_Y}}}$$

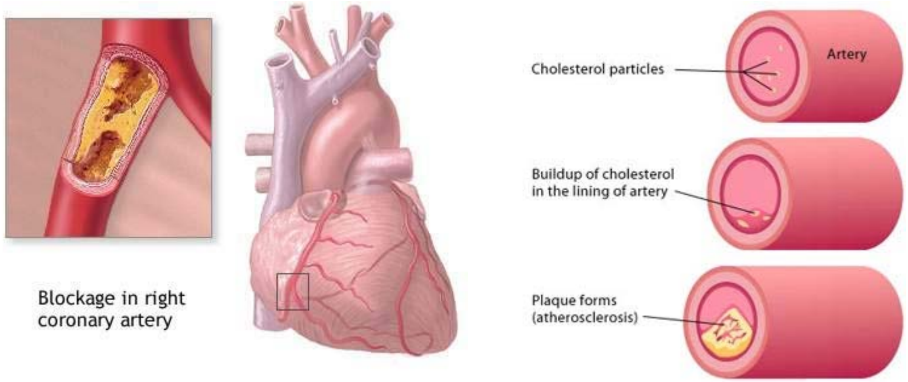
Compute the ***P*-value** :

Null Hypothesis	Alternative Hypothesis	*P* - value
$H_0 : \mu_X - \mu_Y \le \Delta_0$	$H_1 : \mu_X - \mu_Y > \Delta_0$	Area to the right of *z*
$H_0 : \mu_X - \mu_Y \ge \Delta_0$	$H_1 : \mu_X - \mu_Y < \Delta_0$	Area to the left of *z*
$H_0 : \mu_X - \mu_Y = \Delta_0$	$H_1 : \mu_X - \mu_Y \ne \Delta_0$	Sum of the areas in the tails cut off by z and -z

Example 6 (Health Risks)

In a test to compare the effectiveness of two drugs designed to lower cholesterol levels, 75 randomly selected patients were given drug A and 100 randomly selected patients were given drug B. Those given drug A reduced their cholesterol levels by an average of 40 with a standard deviation of 12, and those given drug B reduced their level by an average of 42 with a standard deviation of 15. The units are milligrams of cholesterol per deciliter of blood serum.

Can you conclude that the mean reduction using drug B is greater than that of drug A?

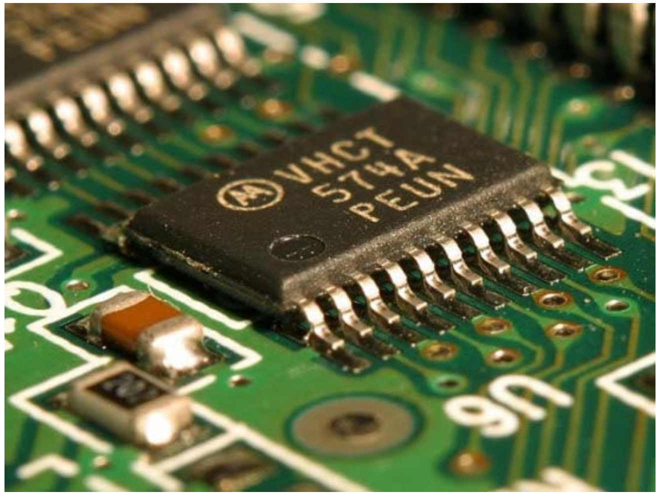


Solution:

Example 7 (New Product Risks)

50 specimens of a new computer chip were tested for speed in a certain application, along with 50 specimens of chips with the old design. The average speed, in MHz, for the new chips was 495.6, and standard deviation was 19.4. The average speed for the old chips was 481.2, and standard deviation was 14.3.

Can you conclude the mean speed for the new chips is greater than that of the old chips?



Solution:

Determine if the spreads of the two sample sets are equal by comparing their standard deviations.

- **Pooled test if $S_{large} / S_{small} < 2$ - the spreads are recognized to be equal.**
- **Un-pooled test if $S_{large} / S_{small} \geq 2$ - the spreads are not equal.**

Independent samples: Small-Sample Tests for the Difference Between Two Means – Pooled tests

$S_{large} / S_{small} < 2$

Compute the test statistic: $$t = \frac{\left(\bar{X} - \bar{Y}\right) - \Delta_0}{S_p \sqrt{1/n_X + 1/n_Y}}$$

Where $$S_p = \sqrt{\frac{(n_X - 1)S_X^2 + (n_Y - 1)S_Y^2}{n_X + n_Y - 2}}$$

The degrees of freedom: $n_X + n_Y - 2$

Compute the ***P*-value** :

Null Hypothesis	Alternative Hypothesis	*P* - value
$H_0 : \mu_X - \mu_Y \leq \Delta_0$	$H_1 : \mu_X - \mu_Y > \Delta_0$	Area to the right of *t*
$H_0 : \mu_X - \mu_Y \geq \Delta_0$	$H_1 : \mu_X - \mu_Y < \Delta_0$	Area to the left of *t*
$H_0 : \mu_X - \mu_Y = \Delta_0$	$H_1 : \mu_X - \mu_Y \neq \Delta_0$	Sum of the areas in the tails cut off by *t* and -*t*

Example 8 (Financial Risks, Part 1)

It is thought that a new process for producing a certain chemical may be cheaper than the currently used process. Two process were run 3 and 6 times respectively, and the cost of producing 100 L of the chemical was determined each time. The results, in dollars, were as follows:

	1	2	3	4	5	6
New Process	51	52	55			
Old Process	50	54	59	56	50	58

Can you conclude the mean cost of the new method is less than that of the old method?

Solution:

Independent samples: Small-Sample Tests for the Difference Between Two Means – Un-pooled tests

$S_{large} / S_{small} \geq 2$

Compute the test statistic: $$t = \frac{\left(\bar{X} - \bar{Y}\right) - \Delta_0}{\sqrt{s_X^2 / n_X + s_Y^2 / n_Y}}$$

The degrees of freedom

$$\nu = \frac{\left[s_X^2 / n_X + s_Y^2 / n_Y \right]^2}{\left[\left(s_X^2 / n_X \right)^2 / \left(n_X - 1 \right) \right] + \left[\left(s_Y^2 / n_Y \right)^2 / \left(n_Y - 1 \right) \right]}$$

Note: the value of the degrees of freedom should be rounded down to the nearest integer

Compute the ***P*-value** :

Null Hypothesis	**Alternative Hypothesis**	***P* - value**
$H_0 : \mu_X - \mu_Y \leq \Delta_0$	$H_1 : \mu_X - \mu_Y > \Delta_0$	Area to the right of *t*
$H_0 : \mu_X - \mu_Y \geq \Delta_0$	$H_1 : \mu_X - \mu_Y < \Delta_0$	Area to the left of *t*
$H_0 : \mu_X - \mu_Y = \Delta_0$	$H_1 : \mu_X - \mu_Y \neq \Delta_0$	Sum of the areas in the tails cut off by *t* and -*t*

Example 9 (Financial Risks, Part 2)

Another new process for producing a certain chemical may be cheaper than the currently used process. Two process were run 3 and 6 times respectively, and the cost of producing 100 L of the chemical was determined each time. The results, in dollars, were as follows:

	1	2	3	4	5	6
New Process	**54**	52	**54**			
Old Process	53	54	59	58	50	58

Can you conclude the mean cost of the new method is less than that of the old method?

Solution:

Dependent samples: Tests with Paired Data

Let (X_1, Y_1), …., (X_n, Y_n) be a sample of ordered pairs whose differences D_1, …., D_n are a sample from normal population with mean μ_D and standard deviation s_D.

Sample is small: $t = \dfrac{\overline{D} - \mu_0}{s_D / \sqrt{n}}$ Sample is large: $z = \dfrac{\overline{D} - \mu_0}{s_D / \sqrt{n}}$

Compute the ***P*-value** :

Null Hypothesis	Alternative Hypothesis	*P* - value
$H_0 : \overline{D} \leq \mu_0$	$H_1 : \overline{D} > \mu_0$	Area to the right of *t or z*
$H_0 : \overline{D} \geq \mu_0$	$H_1 : \overline{D} < \mu_0$	Area to the left of *t or z*
$H_0 : \overline{D} = \mu_0$	$H_1 : \overline{D} \neq \mu_0$	Sum of the areas in the tails cut off by *t* (z) and –*t (z)*

Dependent samples: Tests with Paired Data

Particulate matter emissions from automobiles are a serious environmental concern. Eight vehicles were chosen at random from a fleet, and their emissions were measured under both highway driving and stop-and-go driving conditions (in mg per gallon of fuel) as follows

Vehicle	1	2	3	4	5	6	7	8
Stop-and-go	1500	870	1120	1250	3460	1110	1120	880
Highway	941	456	893	1060	3107	1339	1346	644
Difference	559	414	227	190	353	-229	-226	236

Can we conclude that the mean level of emissions is less for highway driving than for stop-and-go driving?

Hypothesis testing with Paired Data

Solution:

Example (Quality Risks)

The compressive strength, in kilopascals, was measured for each of five concrete blocks both three and six days after pouring. The data are presented in the following table.

	1	2	3	4	5
After 3 days	1341	1316	1352	1355	1327
After 6 days	1376	1373	1366	1384	1358

Can you conclude that the mean strength after six days is greater than the mean strength after three days?

Solution: