SHINGLES ANALYSIS A & B

PROBLEM:

An important quality characteristic used by the manufacturers of ABC asphalt shingles is the amount of moisture the shingles contain when they are packaged. Customers may feel that they have purchased a product lacking in quality if they find moisture and wet shingles inside the packaging.   In some cases, excessive moisture can cause the granules attached to the shingles for texture and colouring purposes to fall off the shingles resulting in appearance problems. To monitor the amount of moisture present, the company conducts moisture tests. A shingle is weighed and then dried. The shingle is then reweighed, and based on the amount of moisture taken out of the product, the pounds of moisture per 100 square feet is calculated. The company would like to show that the mean moisture content is less than 0.35 pound per 100 square feet.

The file (A & B shingles.csv) includes 36 measurements (in pounds per 100 square feet) for A shingles and 31 for B shingles.

NESSCARY MODLUES

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
%matplotlib inline
import seaborn as sns
import scipy.stats as stats
from   scipy.stats              import ttest_1samp, ttest_ind
import statsmodels.stats.api as sm
```

EDA

We have two variable A and B

NULL VALUES are present in B

```
A    0
B    5
```

FIVE POINT SUMMARY

	count	mean	std	min	25%	50%	75%	max
A	36.0	0.316667	0.135731	0.13	0.2075	0.29	0.3925	0.72
B	31.0	0.273548	0.137296	0.10	0.1600	0.23	0.4000	0.58

3.1 Do you think there is evidence that means moisture contents in both types of shingles are within the permissible limits? State your conclusions clearly showing all steps.

SOLUTION:

In this problem we have provided with two independent samples of shingles A and B population standard deviation is unknown and hence we can't perform z test. So we have to go with t-test.

Since we have to find the mean moisture level is less than the permissible limit for the both samples we have perform one sample t-test for sample A and sample B.

SAMPLE A

STEP 1:

DEFINE NULL AND ALTERNATE HYPOTHESIS

The null hypothesis states that the moisture content of sample A is greater or than equal to the permissible limit, $\mu \geq 0.35$

The alternative hypothesis states that the moisture content of sample A is less than permissible limit, $\mu < 0.35$

$H0: \mu \geq 0.35$

$H: \mu < 0.35$

STEP 2:

DECIDE THE SIGNIFICANCE LIMIT

Since alpha value is not given in the question we assume it has alpha $= 0.05$

STEP 3

IDENTIFY THE TEST STATISTIC

We have sample A and we do not know the population standard deviation. Sample size n=36. We use the t distribution and the $tSTAT$ test statistic for one sample t-test.

STEP 4:

CALCULATE THE P - VALUE AND TEST STATISTIC

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

Xbar = 0.316667

S = 0.135731

N = 36

Mu = 0.35

Tstat = -1.4735

(P Value/2) = 0.0747

STEP 5:

DECIDE TO REJECT OR ACCEPT NULL HYPOTHESIS

Since tstat > p_value, we fail to reject the null hypothesis

We conclude that the moisture content is greater than permissible limit in sample A.

SAMPLE B

STEP 1:

DEFINE NULL AND ALTERNATE HYPOTHESIS

The null hypothesis states that the moisture content of sample B is greater or than equal to the permissible limit, $\mu \geq 0.35$

The alternative hypothesis states that the moisture content of sample B is less than permissible limit, $\mu < 0.35$

$H0: \ \mu \geq 0.35$

$H: \ \mu < 0.35$

STEP 2:

DECIDE THE SIGNIFICANCE LIMIT

Since alpha value is not given in the question we assume it has alpha $= 0.05$

STEP 3

IDENTIFY THE TEST STATISTIC

We have sample A and we do not know the population standard deviation. Sample size n=31. We use the t distribution and the $tSTAT$ test statistic for one sample t-test.

STEP 4:

CALCULATE THE P - VALUE AND TEST STATISTIC

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

Xbar = 0.2735

S = 0.1372

N = 31

Mu = 0.35

Tstat = -3.1003

P Value = 0.0020

STEP 5:

DECIDE TO REJECT OR ACCEPT NULL HYPOTHESIS

Since tstat < p_value, we reject the null hypothesis

We conclude that the moisture content is less than permissible limit in sample B

3.2 Do you think that the population mean for shingles A and B are equal? Form the hypothesis and conduct the test of the hypothesis. What assumption do you need to check before the test for equality of means is performed?

STEP 1

DEFINE NULL AND ALTERNATIVE HYPOTHESIS

In testing whether the mean for shingles A and Shingles B are the same, the null hypothesis states that the mean of shingle A to mean of shingle B are the same, μ{A} equals μ{B}. The alternative hypothesis states that the mean are different, $\mu {A}$ is not equal to μ{B}

STEP 2:

DECIDE THE SIGNIFICANCE LIMIT

Since alpha value is not given in the question we assume it has $\mathrm{alpha} = 0.05$

STEP 3

IDENTIFY THE TEST STATISTIC

We have two samples and we do not know the population standard deviation.

Sample sizes for both samples are not the same.

The sample size is, n > 30. So we use the t distribution and the $tSTAT$ test statistic for two sample test.

Two tail test

STEP 4:

CALCULATE THE P - VALUE AND TEST STATISTIC

$t = (M_1 - M_2)/\sqrt{(s^2_{M_1} + s^2_{M_2})}$

CALCULATION:

$N_{1=}$ 36 N_2 =31

M_1= 0.32 M_2 = 0.27

S^2_1 = 0.02 S^2_1 = 0.02

DF_1 = 35 DF2 = 30

Tstat 1.2896282719661123

P Value 0.2017496571835306

STEP 5

DECIDE TO REJECT OR ACCEPT THE NULL HYPOTHESIS

Since tstat > p_value, we fail to reject the null hypothesis

We conclude that mean for shingles A and singles B are not the same