**Practical 6**

**STUDENT TEST (t – DISTRIBUTION)**

When we do not know the population variance or standard deviation we rely of the sample variance and standard deviation. We can use the sample variance (S2) as a best point estimator (approximation) for population variance (∂2), but In this case the distribution will not follow the standard normal distribution (Z-distribution) but the t-distribution.

t-distribution curve is characterized with the following:

1. It has a mean of zero.
2. It is symmetrical around its mean.
3. Its range lie between – α & + α.
4. The quantity (n-1) which is called the degree of freedom (d.f) is used in computing the sample variance.
5. Compared to normal distribution; it is of lower peak & higher tails. This is because the variability is dependent upon S instead of δ and since the variability within the sample is larger than that within the population, then S is usually > than δ.
6. It approaches normal distribution as (n-1) approaches infinity.
7. As the sample size increases and approaches the normal distribution (this is at n > 200), then we will shift from t-distribution to Z-distribution.

**CHOOSING BETWEEN Z & t DISTRIBUTIONS:**

the variability of the sample (S2 &S) approaches that of the population as the sample size increases so the question will be when we can use Z instead of t-distribution (i.e. when to state that S ≈ δ):

* d.f > 200 → use Z test
* d.f 61-200 → use Z or t test
* d.f 31-60 → t test is preferred on Z
* d.f ≤ 30 → we have to use t test

So when ever d.f > 30 we can use Z test & S instead of δ.

**Applications of the t-test**

T-test can be applied for the following situations: (Not much different than of Z- test)

**1-** Is the sample mean differs significantly from the population mean? [Small sample size (n≤30), and the population variance (∂2) is a known].

As in standard normal distribution in which we have the Z-table, here we have the t-table which depend on the **df= (n-1)** → Row of the table & ⍺ (probability of error) = t1 - ⍺/2 → column of the table.

Notes about t-table:

A) The increase in the value of degree of freedom (d.f)

1. From 1 to 30 is by 1 (i.e. 1, 2, 3...).
2. From 30 to 50 is by 5 (i.e. 30, 35, 40…).
3. From 50 to 100 is by 10 (i.e. 50, 60, 70…).
4. From 100 to 200 is by 20 (i.e. 120,140,160…).

B) If we do not find the value of d.f that we want, we choose the nearest one, e.g. 41→ 40, 148→ 140, 96→100, and for 150 either 140 or 160.

C) The table stops at 200 and shifts to ∞, as here we will shift to Z table & Z- distribution.

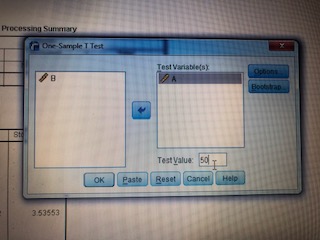
D) To find t-value we use t-table, and z value from z-table.

E) The tabulated value for t depends upon 2 factors; d.f (i.e. sample size & the probability of error (α) → d.ft(1-∞/2), e.g. for n =16, & α=0.1 ; so d.f.=n-1=15, 1-∞/2= 1-(0.1/2)= 1- 0.05= 0.95 then we go to t-table and look where the row of d.f (15) cross with the column of 1- ∞/2 (0.95) to get the tabulated value, so d.ft(1-∞/2) = 15t0.975 = 2.2622.

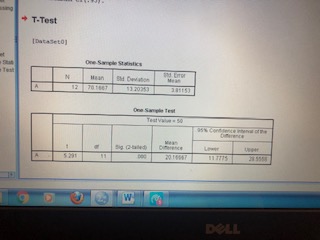
**2-** Comparing the significant difference between two sample means. [Small sample size (n≤30) and the population variance (∂2) is unknown]. We use the following formula:

Where; d.f = n1 + n2 - 2

**Analyze \_\_\_\_ Compare Means \_\_\_\_ One sample t test \_\_\_Select Variable \_\_\_\_ Set Value we want to compare with it \_\_\_\_\_\_\_\_\_\_\_\_ok**

****

**The result will be :**

****

**3- Pairing:** for mean difference (μd) of paired observations (2 dependant samples): Many studies are designed to produce observation in pair e.g., single individual has pair of reading (before & after), for example measurement of BP before and after treatment Or when the same volunteers or participants pass through 2 different situations (each one has 2 readings e.g. as for 2 drugs, 2 different doses for the same drug, drug and placebo, or rest & exhaustion

**EX:** In a study of bran in the treatment of diverticulosis. The clinician wonders whether transit time would be shorter if bran is given in the same dosage in three meals during the day (treatment A) or in one meal (treatment B). A random sample of patients with disease of comparable severity and aged 20-44 is chosen and the two treatments administered on two successive occasions, the results are shown in table below:

|  |  |  |
| --- | --- | --- |
| **Patient** | **Treatment times** | |
| **Treatment A** | **Treatment B** |
| **1** | **63** | **55** |
| **2** | **54** | **62** |
| **3** | **79** | **108** |
| **4** | **68** | **77** |
| **5** | **87** | **83** |
| **6** | **84** | **78** |
| **7** | **92** | **79** |
| **8** | **57** | **94** |
| **9** | **66** | **69** |
| **10** | **53** | **66** |
| **11** | **76** | **72** |
| **12** | **63** | **77** |

Analyze \_\_\_\_\_\_\_\_\_\_\_\_\_\_Compare Meanes\_\_\_\_\_\_\_\_\_\_\_\_\_ Paried t test \_\_\_\_Select variable (A)\_\_\_\_\_then select variable (B) \_\_\_\_\_\_\_\_\_\_Ok

