Statistics for Business

Testing about the population mean: σ unknown

Author : Josemari Sarasola

Gizapedia

gizapedia.hirusta.io

Testing population means: σ unknown

HELLO! Take a look at slide no. 12 about Parametric Testing. In the previous task you learned testing about the μ population mean, assuming that σ, the standard deviation of the population, in known. But in practice σ is usually unknown.

When *sigma* is unknown, we have to estimate it from data, with this formula:

$$\hat{s} = \sqrt{\frac{\sum(x_i - \overline{x})^2}{n-1}}$$

For example if data are 3 and 7 (so arithmetic mean is 5), we calculate this formula in this way:

$$\hat{s} = \sqrt{\frac{(3-5)^2 + (7-5)^2}{2-1}} = 2.82$$

That result is called the *sample* standard deviation (and its squared value, the sample variance). Sometimes you will see another formula, that divides the sum by n instead of by $n-1$. Divided by n, the result is the *population* standard deviation (and its square the population variance). We can prove that dividing by $n-1$ is a better estimate of the real *sigma* (or its square σ^2, the variance). That's because we will divide by $n-1$.

So not knowing σ is not a problem: we estimate it through $\hat{s}$.

But when we estimate σ the sample distributions change as you can see in slide no. 12. In those cases we have two different situations:

1. the population is not normally distributed
2. the population is normally distributed

(1) WHEN POPULATION IS NOT NORMALLY DISTRIBUTED (OR WE CANNOT ASSUME THAT), WE MUST HAVE A BIG SAMPLE TO GET RESULTS.
If that condition is held, the sampling distribution of AM is normal. as you can see on the slide, just like in the previous lesson but taking $\hat{s}$ instead of σ.

Now solve this problem: Usually average grade among students is 6 and we cannot assume that they distribute normally. Due to covid19 emergency, the exam was online and the students took these grades:
7 8 9 6 7 8 7 5 5 6 7 8 8 8 9 9 8 7 6 3 4 7 6 5 7 8 8 8 9 7 6 5 6 7 8 7 8 9 7 9 6 7 7 8 9 9 7 8 9 7 6 5
With a significance level of 10%, can we claim that the mean grade has increased (look at this question to decide is one-sided or tow-sided)? Hint: try to calculate the mean and the sample standard deviation in a compact way to simplify calculations.

(2) WHEN POPULATION IS NORMAL, THE SAMPLING DISTRIBUTION IS A STUDENT T, as you can see on the slide. But what's that? Well, it's a symmetric distributions around 0, very similar to the z standard normal, but with slightly thicker extremes. It has one parameter, named no of degress of freedom. Take a look at 8-9 slides. Values of t distributions with a given no of degrees of freedom for given probabilities are tabulated (take the tables and verify the examples given for the slides). As you can see, the tables gives values for given probabilities. So, it doesn't calculate exact probabilities like the normal table. Hence, when we use this distribution for testing we cannot calculate by hand exact p-values so WE ARE FORCED TO DECIDE USING THE CRITICAL VALUE. Another remark: when degrees of freedom are bigger than 30, we can use normal distribution directly.

Now, if you look to slide no 7. you will see that the sampling distribution is not about the arithmetic mean but about the arithmetic mean standardized with the sample deviation we have calculated from data. This is very important. In these problems, we must always calculate the Am, and then standardize it.

The process is:
(1) Set the null hypothesis about μ
(2) One sided, two-sided?
(3) Set critical value(s) for t and name it t* (sample mean standardized, as on slide 7).
(4) Calculate your own t.
(5) Compare t to t* value(s) and decide about Ho.

Look at this video for an example (you can find more videos searching for "one mean t test"):
https://www.youtube.com/watch?v=vEG$_M OnyMdE$

Remark: here the man sets an alternative hypothesis, added to Ho. Don't care about that. Do it by our usual manner: if $H_0 : \mu = \mu_0$ we have a two-sided test, $H_0 : \mu \geq \mu_0$ we reject on the lower side (critical t value will negative), and if $H_0 : \mu \leq \mu_0$ we reject on the upper side.

Now, solve 152 (only (a) section), 157 (complete) and 158 (only (a) section)

THE 15% EXAM BEFORE MAY 21, WILL BE ABOUT THIS T TEST.