A UNIFIED APPROACH TO TESTING HYPOTHESES ABOUT PARAMETERS OF A NORMAL POPULATION

Nikolai Kolev and Delhi Paiva
University of São Paulo, Brazil
nkolev@ime.usp

We discuss our practice related to classical hypothesis testing about unknown parameters of a normal population offered to undergraduates in the University of São Paulo. We consider the tests for the population mean and variance when the sample size is "large" and "small" as well as the well-known tests comparing the means and variances for independent samples. We suggest an algorithmic approach, which our students appreciate.

INTRODUCTION

Any statistical textbook contains a chapter introducing the basic concepts of testing hypotheses. Depending on the aim of the textbook such a chapter is followed by a description of the corresponding tests about parameters of a normal population. Especially when statistical applications are oriented to Engineering, Economics, Business, Biology, Geography, Psychology, etc, one can find a long list of possible cases (classical and very particular) treating "large" and "small," independent, repeated (paired) and "dependent" samples.

When the first named author tried to present the variety of the types of hypotheses about parameters of a normal population, he lost himself: imagine how the students fared. This was an indication that the topic has to be offered to the students using an algorithmic approach and such an exposition can be found in Kolev (1994).

In this paper we outline our methodology. The exposition is organized as follows. In Section 2 we briefly introduce preliminary concepts and give the algorithmic representation of the material. We finish with several conclusions and related confusions (open questions).

TESTS ABOUT PARAMETERS OF A NORMAL POPULATION

An observation, before it is actually taken (from a population), is modeled as a random variable X with a given distribution. Any numerical feature of a population distribution is called a parameter. Statistical inference deals with drawing conclusions about population parameters from the sample data. Frequently, the objective of an investigation is to decide which of two contradictory claims about a parameter, about the values of several parameters, or about the form of the entire probability distribution, is correct, and the corresponding methods are called hypothesis testing. In general, a statistical hypothesis is a statement about the population. If our population can be identified by a parameter, estimating its numerical value is not of direct interest, nevertheless an estimator may be useful in evaluating the validity of a conjecture.

The null hypothesis, denoted by H_0, is the claim that is initially assumed to be true. The alternative hypothesis, denoted by H_1, is the assertion that is contradictory to H_0. A test statistic is used to measure the difference between the data and what is expected on the null hypothesis (i.e., on "prior belief" claim). The test statistic is a random variable whose realization serves as a base to determine the action (decision) of rejecting of H_0 (in favour of H_1). Therefore, "the null hypothesis says that the sample difference is just due to chance; the alternative hypothesis says that it points to a real difference," (e.g., Freedman *et al*., 1998, p. 479).

In this note the population of interest is normal. We will consider two situations in which the observations form: (i) a random sample from a normal population; (ii) two independent samples from two normal populations. In formulation of the null hypotheses we use a notation "≈" which means one of "≤," "=" or "≥." Let us note that for large samples, the procedures discussed are no longer heavily dependent on the assumption of a normal population.

Testing Hypotheses about Mean and Variance

Let X_1, $X_{2, \ldots,}$ X_n constitute a random sample from a normal distribution with parameters μ and σ^2, to be denoted by $N(\mu, \sigma^2)$. Consequently, the sampling mean $\overline{X}$ is normal with distribution $N(\mu, \sigma^2/n)$.

If our interest is on testing H_0: $\mu \approx \mu_0$, (μ_0 is a specified number) when σ is *known*, then (by virtue of the central limit theorem) the standardized variable

$$Z = \frac{\sqrt{n}(\overline{X} - \mu_0)}{\sigma} \tag{1}$$

has the $N(0,1)$ distribution for *arbitrary* sample size n assuming that H_0 is true. The above procedure is graphically presented on the left-most side in the algorithm shown in Figure 1.

If the parameter σ is *unknown*, then it can be replaced in (1) by the sample standard deviation s_X without affecting the distribution of the corresponding statistic for "large" samples (as a rule, we assume $n \geq 30$).

For "small" samples, i.e. $n < 30$, the statistic

$$T = \frac{\sqrt{n}(\overline{X} - \mu_0)}{s_X}$$

can be approximated by a t distribution with n-1 degrees of freedom, assuming that H_0 is true.

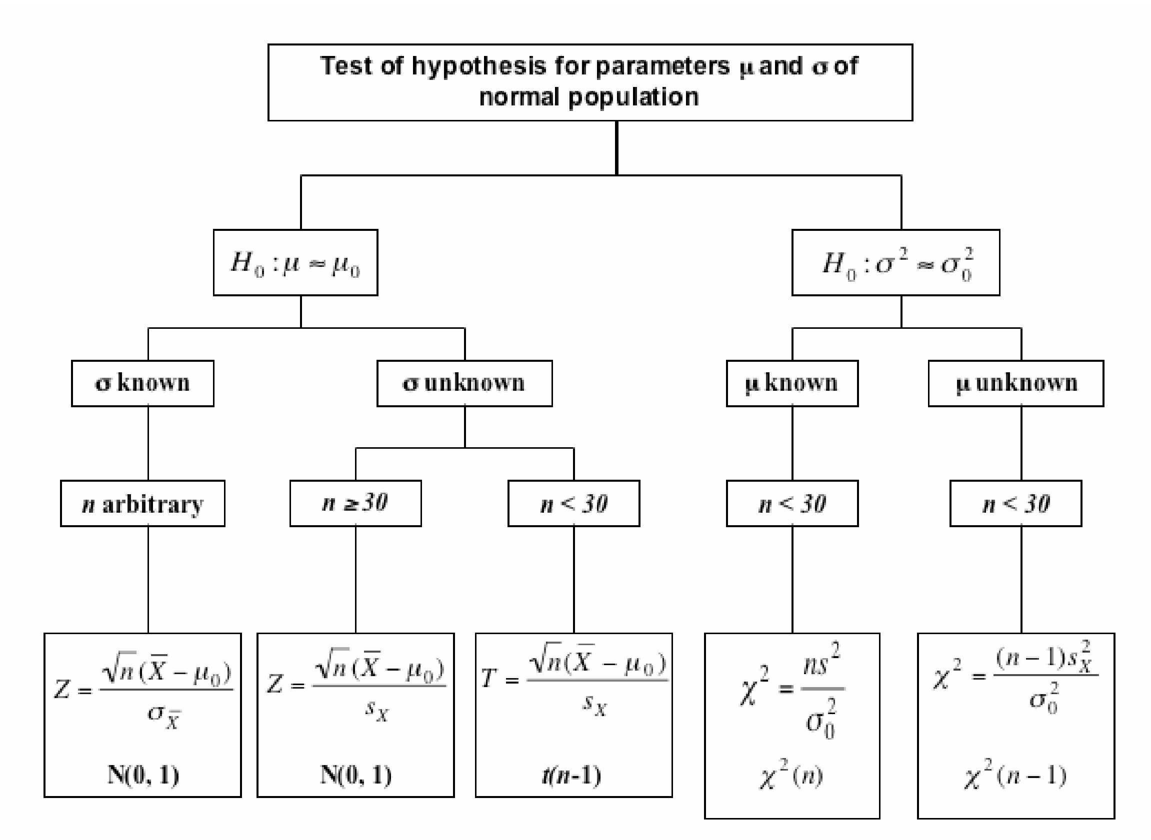

Figure 1

If one needs to test H_0: $\sigma^2 \approx \sigma_0^2$, ($\sigma_0$ is a specified number) when μ is *unknown* and $n <$ 30, then the statistic $(n\text{-}1)s_X^2/\sigma_0^2$ can be approximated by a χ^2 distribution with (n-1) degrees of freedom, i.e., by $\chi^2(n\text{-}1)$. Under the unrealistic assumption that μ is *known*, the last statistic transforms to ns^2/σ_0^2 with distribution $\chi^2(n)$, where $s^2 = (n\text{-}1)s_X^2/n + (\overline{X} - \mu)$. These hypotheses are presented on the right part of the algorithm displayed in Figure 1.

Following the edges on Figure 1, the student easily could find the statistic needed (given in the boxes of the last line along with their asymptotic distribution) to test the corresponding hypothesis, taking into account the available data (and therefore knowing if the sample size is "large" or "small") and of course, the information about the second parameter (is it known or unknown).

Comparing Means and Variances

Let $X_1, X_2, \ldots X_{n1}$ and $Y_1, Y_2, \ldots Y_{n2}$ be random samples from $N(\mu_1,\sigma_1^2)$ and $N(\mu_2,\sigma_2^2)$ distributions, correspondingly. Assume that both samples are independent of one another. The natural estimator of μ_1-μ_2 is $\overline{X}-\overline{Y}$ which is normal with distribution $N(\mu_1-\mu_2,\ \sigma_1^2/n+\sigma_2^2/n)$. Let s_X and s_Y be the sample standard deviations, respectively.

Our interest is to test H_0: μ_1-$\mu_2 \approx \Delta_0$, (Δ_0 is a specified number) when σ_1 and σ_2 are both *known* or *unknown*. Depending on sample sizes n_1 and n_2, the corresponding test statistics are given in the boxes on the last line of Figure 2, where one can find their asymptotic distribution under condition that H_0 is true.

If σ_1 and σ_2 are both *unknown* but *equal*, the numerator of the test statistic $\hat{\sigma}_{\overline{X}-\overline{Y}}$ for small sample sizes is given by

$$\hat{\sigma}_{\overline{X}-\overline{Y}} = [((n_1+ n_2)/n_1n_2)((n_1-1)s_X^2 + (n_2-1)s_Y^2) /(n_1+ n_2)]^{1/2}.$$

If σ_1 and σ_2 are both *unknown* but *different* (which is more realistic), the degrees of freedom ν of the corresponding t distribution is equal to the integer part of the quantity

$$[(s_X^2/ n_1 + s_Y^2/ n_2)^2] /[(s_X^4\, n_1^{-2}/(n_1-1)) + (s_Y^4\, n_2^{-2})/(n_2-1))].$$

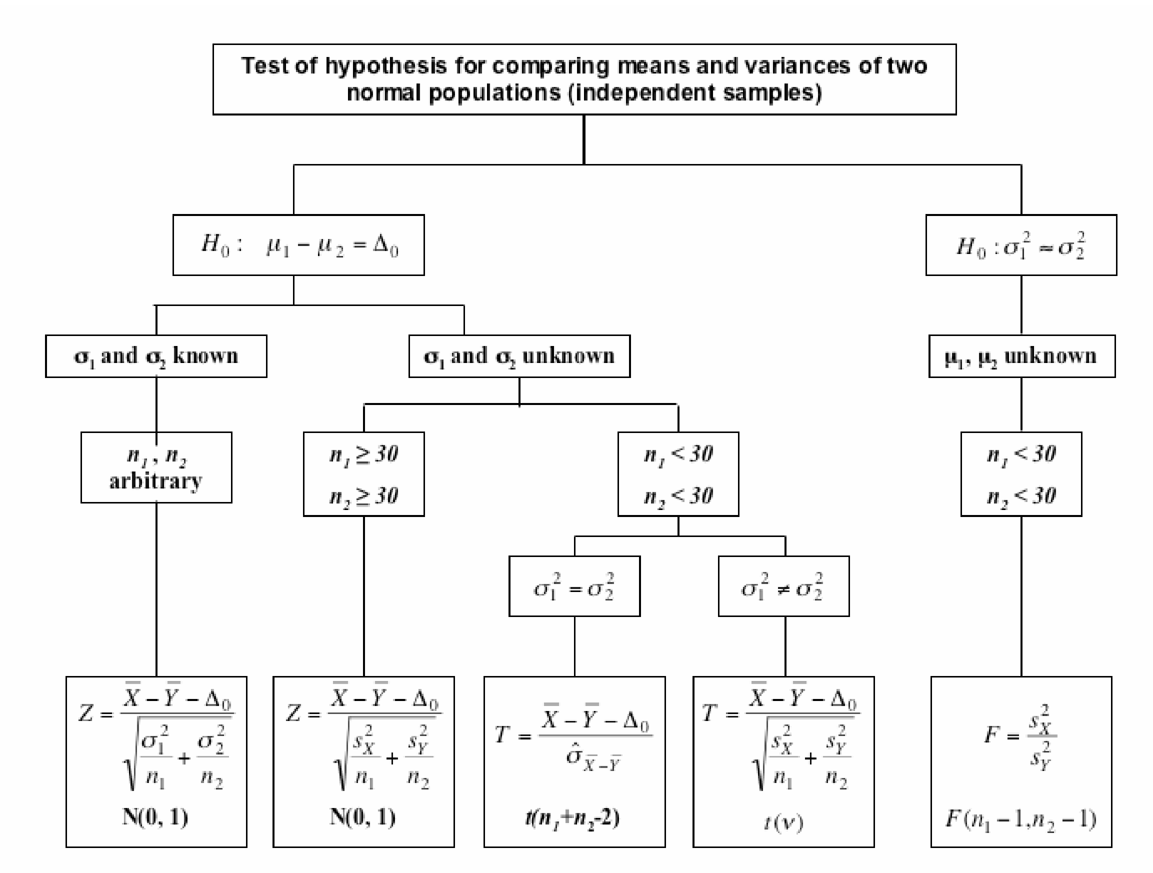

Figure 2

For small samples and unknown μ_1 and μ_2 the test statistic for testing H_0: $\sigma_1^2 \approx \sigma_2^2$ can be approximated by F distribution with n_1-1 and n_2-1 degrees of freedom, as given on the most right part of Figure 2.

Let us note that similar hypotheses can be tested when a sample is from a bivariate normal population, e.g., Kolev (1994). In such a case, the data consist of n independently selected pairs $(X_1,Y_1),(X_2,Y_2),\ldots,(X_n,Y_n)$ from a bivariate normal distribution with parameters μ_1, μ_2, σ_1^2, σ_2^2 and population correlation coefficient ρ_{XY}.

CONFUSIONS AND CONCLUSIONS

There are a lot of confusions concerning the tests related to normal population. At first, nobody can tell us what the "normal" population means. Simply, the real world is bounded, so any application of "normal" theory (allowing infinite realizations) is questionable about a reasonable "approximation" of the quantity of interest. Secondly, the significance level is usually determined intuitively or by an expert, and for the best of our knowledge there is no reliable general rule. There are many other confusions also, but we will stop the possible listing here.

The methodology presented in Section 2 is related to an easy way to apply the right sampling distribution of the test statistic. It reflects our practice for representing the variety of hypothesis regarding parameters of normal population. We do not state that our approach is new or the best one. Probably, it can be found or partially presented in some existing textbook or lecture notes but we did not meet such a methodology.

Because of size limitations we did not discuss fundamental questions like: "What can go wrong if the population distribution is nonnormal?," "What procedures should be used if it is nonnormal?" or "If observations are not independent, is this serious?"

REFERENCES

Freedman, D., Pisani, R. and Purves, R. (1998). *Statistics* (3rd edition). New York: Wiley.

Kolev, N. (1994). *Applied Statistics, Part 1 (With Program PRISTAT 1)*. Sofia: Stopanstvo Publishing House.