A note on Hypothesis Testing

This small note on hypothesis testing is the summary of all testing problems. Following are the working rules for any hypothesis testing problem.

Working rules for Hypothesis testing:

Step 1: State the Null Hypothesis (H_0) and Alternative Hypothesis(H_1)

$\mathbf{H_0}$**:** Select the claim that represents equal (=), greater-equal (≥) or less-equal (≤) relationship with the given population parameter value.

$\mathbf{H_1}$**:** Select the claim that represents not equal (≠), less than (<) or greater than (>) relationship with the given population parameter value.

@remember : H_0 always contains the equal (=) sign.

Step 2: Collect the sample $X_1, X_2, \ldots, X_n$.

Step 3: Calculate the **test** statistic $T_{cal} = f(X_1, X_2, \ldots, X_n)$ depending on the problem statement.

Step 4: Set the level of significance α.

Step 5: Construct Acceptance / Rejection regions depending on α and H_1.

Step 6: Conclusion :

i) Reject the null hypothesis: if T_{cal} falls under critical region,
ii) Do not reject the null hypothesis: if T_{cal} does not fall under critical region.

The main issue for a given hypothesis testing problem is the identification of correct test statistic. Most of the real world problems of hypothesis testing can be broadly classified into 5 major categories and they are:

A) Test of Mean
B) Test of Variance
C) Test of Model validation
D) Test of independence of two variable
E) ANOVA

For a given problem, you have to first identify which of the above category it belongs. Second, from the information given in the problem, you have to identify which of the subcategory (if any) it may belongs. Now, you can perform the hypothesis testing using the appropriate test statistic.

A) Test of Mean

1. Test for specified mean of a single population:

$X_1, \ldots, X_n$ *Is a Sample from a* $\mathcal{N}(\mu, \sigma^2)$

Case 1:

Population σ^2 *Is Known,* $\overline{X} = \sum_{i=1}^{n} X_i/n$

H_0	H_1	Test Statistic TS	Significance Level α Test	p-Value if $TS = t$
$\mu = \mu_0$	$\mu \neq \mu_0$	$\sqrt{n}(\overline{X} - \mu_0)/\sigma$	Reject if $\lvert TS \rvert > z_{\alpha/2}$	$2P\{Z \geq \lvert t \rvert\}$
$\mu \leq \mu_0$	$\mu > \mu_0$	$\sqrt{n}(\overline{X} - \mu_0)/\sigma$	Reject if $TS > z_\alpha$	$P\{Z \geq t\}$
$\mu \geq \mu_0$	$\mu < \mu_0$	$\sqrt{n}(\overline{X} - \mu_0)/\sigma$	Reject if $TS < -z_\alpha$	$P\{Z \leq t\}$

Z is a standard normal random variable.

Case 2:

Population σ^2 *Is Unknown,* $\overline{X}=\sum_{i=1}^{n} X_i/n S^2 = \sum_{i=1}^{n}(X_i - \overline{X})^2/(n-1)$

H_0	H_1	Test Statistic TS	Significance Level α Test	p-Value if $TS = t$
$\mu = \mu_0$	$\mu \neq \mu_0$	$\sqrt{n}(\overline{X} - \mu_0)/S$	Reject if $\lvert TS \rvert > t_{\alpha/2,\, n-1}$	$2P\{T_{n-1} \geq \lvert t \rvert\}$
$\mu \leq \mu_0$	$\mu > \mu_0$	$\sqrt{n}(\overline{X} - \mu_0)/S$	Reject if $TS > t_{\alpha,\, n-1}$	$P\{T_{n-1} \geq t\}$
$\mu \geq \mu_0$	$\mu < \mu_0$	$\sqrt{n}(\overline{X} - \mu_0)/S$	Reject if $TS < -t_{\alpha,\, n-1}$	$P\{T_{n-1} \leq t\}$

T_{n-1} is a t-random variable with $n-1$ degrees of freedom: $P\{T_{n-1} > t_{\alpha,\, n-1}\} = \alpha$.

2. Testing equality of means of two different populations

Case 1: Unpaired t-test

$X_1, \ldots, X_n$ *Is a Sample from a* $\mathcal{N}(\mu_1, \sigma_1^2)$ *Population;* $Y_1, \ldots, Y_m$ *Is a Sample from a* $\mathcal{N}(\mu_2, \sigma_2^2)$ *Population*

The Two Population Samples Are Independent

to Test

$H_0 : \mu_1 = \mu_2$ versus $H_0 : \mu_1 \neq \mu_2$

Assumption	Test Statistic ***TS***	Significance Level ***α*** Test	***p***-Value if ***TS = t***
σ_1, σ_2 known	$\dfrac{\overline{X}-\overline{Y}}{\sqrt{\sigma_1^2/n+\sigma_2^2/m}}$	Reject if $\lvert TS\rvert > z_{\alpha/2}$	$2P\{Z \geq \lvert t\rvert\}$
$\sigma_1 = \sigma_2$	$\dfrac{\overline{X}-\overline{Y}}{\sqrt{\frac{(n-1)S_1^2+(m-1)S_2^2}{n+m-2}}\sqrt{1/n+1/m}}$	Reject if $\lvert TS\rvert > t_{\alpha/2,\, n+m-2}$	$2P\{T_{n+m-2} \geq \lvert t\rvert\}$
n, m large	$\dfrac{\overline{X}-\overline{Y}}{\sqrt{S_1^2/n+S_2^2/m}}$	Reject if $\lvert TS\rvert > z_{\alpha/2}$	$2P\{Z \geq \lvert t\rvert\}$

Case 2 : Paired t-test

- Before-and-after observations on the same subjects (e.g. students' diagnostic test results before and after a particular module or course).

Let x = test score before the module, y = test score after the module

To test the null hypothesis that the true mean difference is zero, the procedure is as follows:

1. Calculate the difference ($d_i = y_i - x_i$) between the two observations on each pair, making sure you distinguish between positive and negative differences.
2. Calculate the mean difference, $\bar{d}$.
3. Calculate the standard deviation of the differences, s_d, and use this to calculate the standard error of the mean difference, $SE(\bar{d}) = \frac{s_d}{\sqrt{n}}$
4. Calculate the t-statistic, which is given by $T = \frac{\bar{d}}{SE(\bar{d})}$. Under the null hypothesis, this statistic follows a t-distribution with $n-1$ degrees of freedom.
5. Use tables of the t-distribution to compare your value for T to the t_{n-1} distribution. This will give the p-value for the paired t-test.

B) Test of Variance

1. Test for specified variance of a normal population

Let $X_1, \ldots, X_n$ denote a sample from a normal population having unknown mean μ and unknown variance σ^2, and suppose we desire to test the hypothesis

$$H_0 : \sigma^2 = \sigma_0^2$$

versus the alternative

$$H_1 : \sigma^2 \neq \sigma_0^2$$

for some specified value σ_0^2.

Test Statistic:

$$\frac{(n-1)S^2}{\sigma_0^2} \sim \chi^2_{n-1}$$

$$\begin{array}{lll} \text{accept} & H_0 & \text{if} \quad \chi^2_{1-\alpha/2,n-1} \leq \dfrac{(n-1)S^2}{\sigma_0^2} \leq \chi^2_{\alpha/2,n-1} \\ \text{reject} & H_0 & \text{otherwise} \end{array}$$

2. Test for equality of variances of two normal populations

Let $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_m$ denote independent samples from two normal populations having respective (unknown) parameters μ_x, σ_x^2 and μ_y, σ_y^2 and consider a test of

$$H_0 : \sigma_x^2 = \sigma_y^2 \qquad \text{versus} \qquad H_1 : \sigma_x^2 \neq \sigma_y^2$$

Test Statistic:

$$S_x^2 / S_y^2 \sim F_{n-1,m-1}$$

$$\begin{array}{lll} \text{accept} & H_0 & \text{if} \quad F_{1-\alpha/2,n-1,m-1} < S_x^2/S_y^2 < F_{\alpha/2,n-1,m-1} \\ \text{reject} & H_0 & \text{otherwise} \end{array}$$

C) Test of Model validation (Goodness of fit)

A chi-squared test can be used to test the hypothesis that observed data follow a particular distribution. The test procedure consists of arranging the n observations in the sample into a frequency table with k classes. The chi-squared statistic is:

$$\chi^2_{data} = \sum \frac{(O_i - E_i)^2}{E_i}$$

Where O_i = observed frequency, and E_i = expected frequency.

The number of degrees of freedom is $k - p - 1$ where p is the number of parameters estimated from the (sample) data used to generate the hypothesised distribution.

Note, Goodness-of-fit hypothesis are always right tailed.

And state the rejection rule.

Reject if $\chi^2_{data} > \chi^2_{critical}$.

D) Test of independence of two variable

A **test of independence** assesses whether unpaired observations on two variables, expressed in a contingency table, are independent of each other

H_0:	The two-way table is independent
H_a:	The two-way table is not independent
Test Statistic:	$T = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{(O_{ij}-E_{ij})^2}{E_{ij}}$

where

r = the number of rows in the contingency table

c = the number of columns in the contingency table

O_{ij} = the observed frequency of the ith row and jth column

E_{ij} = the expected frequency of the ith row and jth column

$= \frac{R_i C_j}{N}$

R_i = the sum of the observed frequencies for row i

C_j = the sum of the observed frequencies for column j

N = the total sample size

Significance Level:	α
Critical Region:	T > CHSPPF(α,(r-1)*(c-1))
	where CHSPPF is the percent point function of the chi-square distribution and (r-1)*(c-1) is the degrees of freedom
Conclusion:	Reject the independence hypothesis if the value of the test statistic is greater than the chi-square value.

F) ANOVA

ANOVA is used to test whether there is significant difference between means of different groups. Suppose there are *m* groups and each group has *n* observations.

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_m$$

$$H_1 : \text{not all the means are equal}$$

Mean of i-th group	$X_{i.} = \sum_{j=1}^{n} X_{ij}/n$
Overall Mean	$X_{..} = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n} X_{ij}}{nm} = \frac{\sum_{i=1}^{m} X_{i.}}{m}$

$$SS_W = \sum_{i=1}^{m}\sum_{j=1}^{n}(X_{ij} - X_i)^2$$

$$SS_b = n\sum_{i=1}^{m}(X_{i.} - X_{..})^2$$

$$TS = \frac{SS_b/(m-1)}{SS_W/(nm-m)}$$

$$\begin{array}{lll} \text{reject} & H_0 & \text{if} \quad \frac{SS_b/(m-1)}{SS_W/(nm-m)} > F_{m-1,nm-m,\alpha} \\ \text{do not reject} & H_0 & \text{otherwise} \end{array}$$