Republic of Iraq Ministry of Higher Education & Research

University of Anbar

College of Education for Pure Sciences

Department of Mathematics

محاضرات الاحصاء ٢

مدرس المادة : الاستاذ المساعد

الدكتور فراس شاكر محمود

TEST ABOUT ONE MEAN

We begin this chapter on tests of statistical hypotheses with an application in which we define many of the terms associated with testing .

Example :- Let X equal the breaking strength of a steel bar . If the bar is manufactured by process I , X is N(50,36) i.e., X is normally distributed with $\mu = 50$ and $\sigma^2 = 36$. It is hoped that if process II (a new process) is used , X will be N(50,36) . Given a large number of steel bars manufactured by process II , how could we test whether the five-unit increase in the mean breaking strength was realized ?

In this problem , we are assuming that X is $N(\mu, 36)$ and μ is equal to 50 or 55 .We want to test the simple null hypothesis $H_o: \mu = 50$ against the simple alternative hypothesis $H_1: \mu = 55$. Note that each of these hypotheses completely specifies the distribution of X. That is , H_0 states that X is N(50,36) and H_1 states that X is N(55,36) . (If the alternative hypothesis had been $H_1: \mu > 50$, it would be a composite hypothesis , because it is composed of all normal distributions with =36 and means greater than 50) In order to test which of the two hypotheses , H_o or H_1 , is true , we shall set up a rule based on the breaking strengths $x_1, x_2, \dots, x_n$ of n bars (the observed values of a random sample of size n from this new normal distribution) . The rule leads to a decision to accept or reject H_o; hence , it is necessary to partition the sample space into two parts-say , C and C' - so that if $(x_1, x_2, \dots, x_n) \in C$, H_o is rejected , and if $(x_1, x_2, \dots, x_n) \in C'$, H_o is accepted (not reject) . The rejection region C for H_o is called the critical region for the test . Often , the partitioning of the sample space is specified in terms of the values of a statistic called the test statistic . In this example , we could let $\bar{x}$ be the test statistic and say , take $C = \{(x_1, x_2, \dots, x_n): \bar{x} \geq 53\}$; that is we will reject H_o if $\bar{x} \geq 53$. If $(x_1, x_2, \dots, x_n) \in C$ when H_o is true , H_0 would be rejected when it is true , a Tape I error . If $(x_1, x_2, \dots, x_n) \in C'$ when H_1 is true , H_o would be accepted (i.e. not rejected) when in fact H_1 is true , a Type II error . The probability of a Type I error is called the **significance level** of the test and is denoted by α .

That is , $\alpha = P[(x_1, x_2, \dots, x_n) \in C;\ H_o]$ is the probability that $(x_1, x_2, \dots, x_n)$ falls into C when H_o is true . The probability of a Type II error is denoted by β ; that is , $\beta = P[(x_1, x_2, \dots, x_n) \in C';\ H_1]$ is the probability of accepting (failing to reject) H_o when it is false.

As an illustration , suppose n=16 bars were tested and $C = \{\bar{x}: \bar{x} \geq 53\}$. Then $\bar{X}$ is N(50,36/16) when H_o is true and is N(55,36/16) when H_1 is true . Thus ,

$$\alpha = P(\bar{X} \geq 53; H_o) = P\left(\frac{\bar{X}-50}{\frac{6}{4}} \geq \frac{53-50}{\frac{6}{4}}\ ; H_o\right)$$

$$= 1 - \emptyset(2) = 0.0228$$

And

$$\beta = P(\bar{X} \geq 53; H_1) = P\left(\frac{\bar{X}-55}{\frac{6}{4}} \geq \frac{53-55}{\frac{6}{4}}\ ; H_1\right)$$

$= \emptyset\left(-\frac{4}{3}\right) = 1 - 0.9087 = 0.0913$.

Figure 8.1-1 shows the graphs of the probability density functions of $\bar{X}$ when H_o and H_1 , respectively , are true . Note that by changing the critical region , C , it is possible to decrease (increase) the size of α but this leads to an increase (decrease) in the size of β . Both α and β can be decreased if the sample size n is increased.

Example :- Assume that the underlying distribution is normal with unknown mean μ but known variance $\sigma^2 = 100$. Say we are testing the simple null hypotheses $H_o: \mu = 60$ against the composite alternative hypotheses $H_1;\ \mu > 60$ with a sample mean $\bar{X}$ based on n=52 observations . Suppose that we obtain the observed sample mean of $\bar{X} = 62.75$. If we compute the probability of obtaining an $\bar{X}$ of that value of 62.75 or greater when $\mu = 60$. then we obtain the P-value associated with $\bar{X} = 62.75$.

$$P - value = P(\bar{X} \geq 62.75\,; \mu = 60)$$

$$= P\left(\frac{\bar{X}-60}{\frac{10}{\sqrt{52}}} \geq \frac{62.75-60}{\frac{10}{\sqrt{52}}}\ ;\ \mu = 60\right)$$

$$= 1 - \emptyset\left(\frac{62.75-60}{\frac{10}{\sqrt{52}}}\right) = 1 - \emptyset(1.983) = 0.0237.$$

If this p-value is small , we tend to reject the hypotheses $H_o: \mu = 60$. For example rejecting $H_o: \mu = 60$ if the p-value is less than or equal to $\alpha = 0.05$ is exactly the same as rejecting H_o if

$\bar{X} \geq 60 + (1.645)\left(\frac{10}{\sqrt{52}}\right) = 62.281$.

Here $\mathrm{P-value} = 0.0237 < \alpha = 0.05$ and $\bar{\mathrm{X}} = 62.75 > 62.281$

To help the reader keep the definition of p-value in mind , we note that it can be thought of as that tail-end probability , under H_o , of the distribution of the statistic (here $\bar{\mathrm{X}}$) beyond the observed value of the statistic .

If the alternative were the two-sided $H_1 : \mu \neq 60$, then the p-value would have been double 0.0237 ; that is , then the p-value =2(0.0237)=0.0474 because we include both tails .

To test $H_0 : \mu = \mu_0$ against one of these three alternative hypotheses , a random sample is take from the distribution and an observed sample mean , $\bar{X}$, that is close to μ_0 supports H_o. The closeness of $\bar{X}\ to\ \mu_o$ is measured in terms of standard deviations of $\bar{X}\ , \alpha/\sqrt{n}$, when α is known , a measure that is sometimes called the standard error of the mean . Thus the test statistic could be defined by

H_o	H_1	$Critical\ Region$
$\mu = \mu_0$	$\mu > \mu_0$	$z \geq z_\alpha\ or\ \bar{X} \leq \mu_0 + z_\alpha \sigma/\sqrt{n}$
$\mu = \mu_0$	$\mu < \mu_0$	$z \leq -z_\alpha\ or\ \bar{X} \leq \mu_0 - z_\alpha \sigma/\sqrt{n}$
$\mu = \mu_0$	$\mu \neq \mu_0$	$\lvert z \rvert \geq z_\alpha/2\ or\ \lvert \bar{X} - \mu_0 \rvert \geq z_\alpha/\sigma/\sqrt{n}$

$$Z = \frac{\bar{X} - \mu_0}{\sqrt{\sigma^2/n}} = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$$

And the critical regions , at a significance level α , for the three respective alternative hypotheses would be $(i) Z \geq Z_\alpha$, $(ii)\ Z \leq -Z_\alpha$, and $(iii) \lvert Z \rvert \geq Z_{\alpha/2}$. In terms of $\bar{X}$, these three critical regions become $(i) \bar{X} \geq \mu_0 + Z_\alpha \left(\frac{\sigma}{\sqrt{n}}\right)$, $(ii) \bar{X} \leq \mu_0 - Z_\alpha \left(\frac{\sigma}{\sqrt{n}}\right)$, and $(iii) \lvert \bar{X} - \mu_0 \rvert \geq Z_{\frac{\alpha}{2}} \left(\frac{\sigma}{\sqrt{n}}\right)$.

The three tests and the distribution is $N(\mu\ , \sigma^2)$ and σ^2 is known.

It is usually the case that the variance σ^2 is not known . Accordingly , we now take a more realistic position and assume that the variance is unknown . Suppose our null hypotheses is $H_o : \mu = \mu_0$ and the two-sided alternative hypotheses is $H_1 : \mu \neq \mu_0$.

Recall from section 7.1 for a random sample $x_1, x_2, \dots, x_n$ taken a normal distribution $N(\mu , \sigma^2)$, a confidence interval for μ is based on

$$T = \frac{\bar{X} - \mu}{\sqrt{s^2/n}} = \frac{\bar{X} - \mu}{s/\sqrt{n}}$$

This suggests that T might be a good statistic to use for the test of $H_o : \mu = \mu_0$ with μ replaced by μ_0 . In addition , it is the natural statistic to use if we replace σ^2/n by its unbiased estimator s^2/n in $(\bar{X} - \mu_0)/\sqrt{\sigma^2/n}$ in Equation 8.1-1 . If $\mu = \mu_0$, we know that T has a t distribution with n-1 degrees of freedom . Thus , with $\mu = \mu_0$,

$$P\left[|T| \geq t_{\frac{\alpha}{2}}(n-1)\right] = P\left[\frac{|\bar{X} - \mu_0|}{\frac{s}{\sqrt{n}}} \geq t_{\frac{\alpha}{2}}(n-1)\right] = \alpha$$

Accordingly , if $\bar{X}$ and s are , respectively , the sample mean and sample standard deviation , then the rule that reject $H_o: \mu = \mu_o$ and accepts $H_1: \mu \neq \mu_o$ if and only if

$$|t| = \frac{|\bar{X} - \mu_0|}{\frac{s}{\sqrt{n}}} \geq t_{\frac{\alpha}{2}}(n-1)$$

Provides a test of this hypotheses with significance level α . Note that this rule is equivalent to rejecting $H_0: \mu = \mu_0$ if not the open $100(1-\alpha)\%$ confidence interval

$$\left(\bar{X} - t_{\frac{\alpha}{2}}(n-1)\left[\frac{s}{\sqrt{n}}\right] , \bar{X} + t_{\frac{\alpha}{2}}(n-1)\left[\frac{s}{\sqrt{n}}\right]\right)$$

The following Table summarizes tests of hypotheses for a single mean . along with the three possible alternative hypotheses , when the underlying distribution is $N(\mu , \sigma^2)$, σ^2 is unknown , $t = \frac{(\bar{X} - \mu_0)}{\frac{s}{\sqrt{n}}}$. and $n \leq 30$. If $n > 30$, we use the following Table for approximate tests , with σ replaced s .

H_o	H_1	$Critical\ Region$
$\mu = \mu_0$	$\mu > \mu_0$	$z \geq z_\alpha\ or\ \bar{X} \leq \mu_0 + z_\alpha \sigma/\sqrt{n}$
$\mu = \mu_0$	$\mu < \mu_0$	$z \leq -z_\alpha\ or\ \bar{X} \leq \mu_0 - z_\alpha \sigma/\sqrt{n}$
$\mu = \mu_0$	$\mu \neq \mu_0$	$\lvert z\rvert \geq z_\alpha/2\ or\ \lvert\bar{X} - \mu_0\rvert \geq z_\alpha/\sigma/\sqrt{n}$

Example : Let X (in millimeters) equal the growth in 15 days of a tumor induced in a mouse . Assume that the distribution of X is $N(\mu\,, \sigma^2)$. We shall test the null hypotheses $H_0: \mu = \mu_0 = 4.0\ mm$ aganst the two-sided alternative hypothesis $H_1: \mu \neq 4.0$ If we use n=9 observations and a significance level of $\alpha = 0.10$, the critical region is

$$|t| = \frac{|\bar{X} - 4.0|}{\frac{s}{\sqrt{9}}} \geq t_{\frac{\alpha}{2}}(8) = 1.860 \ .$$

If we are given that $n = 9\,, \bar{X} = 4.3\,, and\ s = 1.2$ we see that

$$t = \frac{4.3 - 4.0}{\frac{1.2}{\sqrt{9}}} = \frac{0.3}{0.4} = 0.75 \ .$$

Thus ,

$$|t| = |0.75| < 1.860$$

And we accept (do not reject) $H_o: \mu = 4.0$ at the $\alpha = 10\%$ significance level . (See Figure 8.1-3) The p-value is the two-sided probability of $|T| \geq 0.75$, namely .

$$P - \text{value} = \text{P}(|\text{T}| \geq 0.75) = 2\text{P}(\text{T} \geq 0.75)$$

With our t tables with eight degrees of freedom , we cannot find this P-value exactly . It is about 0.50 , because

$$P(|T| \geq 0.706) = 2P(T \geq 0.706) = 0.50$$

Remark : In discussing the test of a statistical hypothesis , the word accept H_0 might better be replaced by do not reject H_o . That is , if , in Example 8.1-3 , $\bar{X}$ is close enough to 4.0 so that we accept $\mu = 4.0$, we do not want that acceptance to imply that μ is actually equal to 4.0 , We want to say that the data do not deviate enough from $\mu = 4.0$ for us to reject that hypothesis : that is , we do not reject $\mu = 4.0$ with these observed data . With this understanding , we sometimes use accept . and sometimes fail to reject or do not reject , the null hypothesis .

Example : In attempting to control the strength of the wastes discharged into a nearby river , a paper firm has taken a number of measures. Members of the firm believe that they have reduced the oxgen-consuming power of their wastes from a previous mean μ of 500 (measured in parts per million of permanganate) . They plan to test $H_o: \mu = 500$ against $H_1: \mu < 500$, using readings taken on n=25 consecutive days. If these 25 values can be treated as a random sample , then the critical region , for a significance level of $\alpha = 0.01$, is

$$t = \frac{\bar{X} - 500}{\frac{s}{\sqrt{25}}} \leq -t_{0.01}(24) = -2.492$$

The observed values of the sample mean and sample standard deviation wew $\bar{X} = 308.8$ and $s = 115.15$. since

$$t = \frac{308.8 - 500}{\frac{115.15}{\sqrt{25}}} = -8.30 < -2.492$$

We clearly reject the null hypothesis and accept $H_1: \mu < 500$. Note , however , that although an improvement has been made , there stil might exist the question of whether the improvement is adequate . The one-sided 99% confidence interval for μ , namely

$$\left[0,308.8 + 2.492\left(\frac{115.25}{\sqrt{25}}\right)\right] = [0,366.191.],$$

Provides an upper bound for μ and may help the company answer this question.