Module 6: Hypothesis Testing

Hypothesis Testing: Z test for single mean

TEST STATISTIC for single mean

Under the NULL Hypothesis Ho that the sample has been drawn from a population with mean μ and variance σ^2, i.e., there is no signilicant difference between the sample mean ($\bar{x}$) and population mean (μ), the test statistic (for large samples), is

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

Z test for single mean

A stenographer claims that she can take dictation at the rate of 120 words per minute. Can we reject her claim on the basis of 100 trials in which she demonstrates a mean of 116 words with a standard deviation of 15 words ? Use 5 per cent level of significance.

Z test for single mean

Solution

H0: Stenographer's claim is true i.e. $\mu = 120$

Ha: Stenographer's claim is not true i.e. $\mu \neq 120$

$\mu =$	120
$n =$	100
$\bar{x} =$	116
$s =$	15

as n>=30 , large sample and for large samples $\sigma = s$

$\sigma = s =$ 15

Test statistic Z= $\dfrac{-4}{1.5}$

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

Test statistic Z= -2.667

$|z| =$ 2.67

Zcritical = 1.96

CRITICAL VALUES OF Z				
	LEVEL OF SIGNIFICANCE (α)			
	1%	2%	5%	10%
Two-tailed test	$\lvert z_{\alpha/2} \rvert$=2.58	$\lvert z_{\alpha/2} \rvert$=2.33	$\lvert z_{\alpha/2} \rvert$=1.96	$\lvert z_{\alpha/2} \rvert$=1.645
Right-tailed test	z_α=2.33	z_α=2.05	z_α=1.645	z_α=1.28
Left-tailed test	$-z_\alpha$=-2.33	$-z_\alpha$=-2.05	$-z_\alpha$=-1.645	$-z_\alpha$=-1.28

As z > zcritical, we reject the Null Hypothesis at 0.05 LOS

Conclusion : The Stenographer's claim that she can dictate at the rate of 120 words per minute is not true.

Z test for single mean

A company manufacturing automobile tyres finds that tyre-life is normally distributed with a mean of 40,000 km and standard deviation of 3,000 km. It is believed that a change in the production process will result in a better product and the company has developed a new tyre. A sample of 100 new tyres has been selected. The company has found that the mean life of these new tyres is 40,900 km. Can it be concluded that the new tyre is significantly better than the old one, using the significance level of 0.01?

Solution: we have to test whether the mean life of new tyres has increased beyond 40,000 km.

H0: Let the mean life of new tyres be $\mu = 40000$ km

H1: Let the mean life of new tyres $\mu > 40000$ km

$\mu =$ 40000

$n =$ 100

$\bar{x} =$ 40900

$\sigma =$ 3000

as n>=30 , large sample we can use z test for single mean

α %	15 %	10 %	5 %	4 %	1%	0.5 %	0.2 %
α	0.15	0.1	0.05	0.04	0.01	0.005	0.002
$-Z_{\alpha/2}$ and $+Z_{\alpha/2}$	-1.44	-1.645	-1.96	-2.06	-2.58	-2.81	-3.08
for T.T.T.	1.44	1.645	1.96	2.06	2.58	2.81	3.08
$-Z_{\alpha}$ for L.O.T.T.	-1.04	-1.28	-1.645	-2.6	-2.33	-2.58	-2.88
Z_{α} for R.O.T.T.	1.04	1.28	1.645	2.6	2.33	2.58	2.88

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

Standard Error = 300

Test statistic Z= $\frac{900}{300}$

Test statistic Z = 3

|z|= 3

Zcritical = 2.33

As zcalc > zcritical, we reject the Null Hypothesis at 0.01 LOS

Conclusion : We cannot conclude that the new tyre is significantly better than the old one, at the significance level of 0.01

Exercise 1

The length of life X of certain computers is approximately <u>normally distributed</u> with mean 800 hours and standard deviation of 40 hours. If a random sample of 30 computers has an average life of 788 hours. Test the null hypothesis (Ho) that $\mu = 800$ hours against the alternative hypothesis $\mu \neq 800$ hours at 0.5%, 1%, 4%, 5%, 10%, and 15% LOS. (Hint: Use *z* test).

Note: (1) n>=30 use *z* test (2) n<30 use *t* test

Solution to Exercise 1

mean (μ) = 800 hours

standard deviation (σ) = 40 hours

sample size (n) = 30

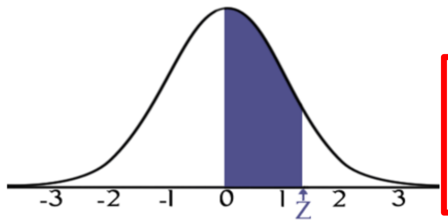
average life ($\bar{x}$) = 788 hours

α %	15 %	10 %	5 %	4 %	1%	0.5 %	0.2 %
α	0.15	0.1	0.05	0.04	0.01	0.005	0.002
$-Z_{\alpha/2}$ and $+Z_{\alpha/2}$ for T.T.T.	− 1.44 1.44	− 1.645 1.645	− 1.96 1.96	− 2.06 2.06	− 2.58 2.58	− 2.81 2.81	− 3.08 3.08
$-Z_{\alpha}$ for L.O.T.T.	− 1.04	− 1.28	− 1.645	− 2.6	− 2.33	− 2.58	− 2.88
Z_{α} for R.O.T.T.	1.04	1.28	1.645	2.6	2.33	2.58	2.88

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

0.5% (-2.81) [Do not reject H_0], 1% (-2.58) [Do not reject H_0], 4% (-2.06) [Do not reject H_0], 5% (-1.96) [Do not reject H_0], 10% (-1.645) [Do not reject H_0], and 15% (-1.44) [Reject H_0] LOS.

$$z_{cal} = -1.643$$



STANDARD NORMAL TABLE (Z)

Entries in the table give the area under the curve between the mean and *z* standard deviations above the mean. For example, for *z* = 1.25 the area under the curve between the mean (0) and *z* is 0.3944.

Z_{table} = 1.2 + 0.05 = 1.25 → 0.3944

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0190	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2969	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3513	0.3554	0.3577	0.3529	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990
3.1	0.4990	0.4991	0.4991	0.4991	0.4992	0.4992	0.4992	0.4992	0.4993	0.4993
3.2	0.4993	0.4993	0.4994	0.4994	0.4994	0.4994	0.4994	0.4995	0.4995	0.4995
3.3	0.4995	0.4995	0.4995	0.4996	0.4996	0.4996	0.4996	0.4996	0.4996	0.4997
3.4	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4998

Exercise 2

Mice with an average lifespan of 32 month will live up to 40 months when fed by a certain nutrious food. If 64 mice fed on this diet have average lifespan of 38 months and standard deviation of 5.8 months, is there any reason to believe that lifespan is <40 months. Use 1% LOS.

α %	15 %	10 %	5 %	4 %	1%	0.5 %	0.2 %
α	0.15	0.1	0.05	0.04	0.01	0.005	0.002
$-Z_{\alpha/2}$ and $+Z_{\alpha/2}$ for T.T.T.	− 1.44 1.44	− 1.645 1.645	− 1.96 1.96	− 2.06 2.06	− 2.58 2.58	− 2.81 2.81	− 3.08 3.08
$-Z_{\alpha}$ for L.O.T.T.	− 1.04	− 1.28	− 1.645	− 2.6	− 2.33	− 2.58	− 2.88
Z_{α} for R.O.T.T.	1.04	1.28	1.645	2.6	2.33	2.58	2.88

$z_{cal} = -2.76$ [Reject H_o]

Exercise 3

A machine runs on an average of 125 hours/year. A random sample of 49 machines has an annual average of 126.9 hours with standard deviation 8.4 hours. Does this suggest to believe that machines are used on the average more than 125 hours annually at 5% LOS.

α %	15 %	10 %	5 %	4 %	1%	0.5 %	0.2 %
α	0.15	0.1	0.05	0.04	0.01	0.005	0.002
$-Z_{\alpha/2}$ and $+Z_{\alpha/2}$ for T.T.T.	−1.44 1.44	−1.645 1.645	−1.96 1.96	−2.06 2.06	−2.58 2.58	−2.81 2.81	−3.08 3.08
$-Z_\alpha$ for L.O.T.T.	−1.04	−1.28	−1.645	−2.6	−2.33	−2.58	−2.88
Z_α for R.O.T.T.	1.04	1.28	1.645	2.6	2.33	2.58	2.88

$z_{cal} = 1.58$ [Do not reject H_o]

Exercise 4

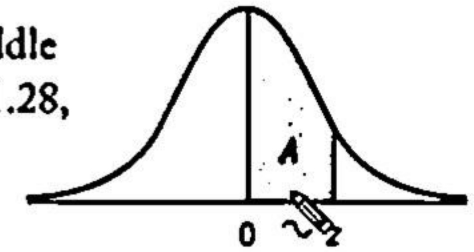
Can it be concluded that the average lifespan of Indian is > 70 years if a random sample of 100 Indians has an average lifespan of 71.8 years with σ of 8.9 years. Use 5% LOS.

$z_{cal} = 2.022$ [Reject H_o]

α %	15 %	10 %	5 %	4 %	1%	0.5 %	0.2 %
α	0.15	0.1	0.05	0.04	0.01	0.005	0.002
$-Z_{\alpha/2}$ and $+Z_{\alpha/2}$ for T.T.T.	− 1.44 1.44	− 1.645 1.645	− 1.96 1.96	− 2.06 2.06	− 2.58 2.58	− 2.81 2.81	− 3.08 3.08
$-Z_{\alpha}$ for L.O.T.T.	− 1.04	− 1.28	− 1.645	− 2.6	− 2.33	− 2.58	− 2.88
Z_{α} for R.O.T.T.	1.04	1.28	1.645	2.6	2.33	2.58	2.88

Normal Curve (z) Table

The normal curve table gives only the percentage of data starting from the middle ($z = 0$), out to whatever z score you look up. For instance, if you look up $z = 1.28$, you get .3997. This means .3997 or 39.97% of the data in the normal curve is found between $z = 0$ and $z = 1.28$.



				Normal						
z	**0.00**	**0.01**	**0.02**	**0.03**	**0.04**	**0.05**	**0.06**	**0.07**	**0.08**	**0.09**
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706

Confidence intervals

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0190	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2969	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3513	0.3554	0.3577	0.3529	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	$\boxed{0.4750}$	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817

- Find $z_{\alpha/2}$ for a 95% confidence interval

Solution

Here are the steps to finding $z_{\alpha/2}$ (When a t-table can't be used)

1. Identify the C-level
2. Find the (C-Level)/2
3. Go to the Z-table, (in the body of the table) look up the number found in step 2
4. $z_{\alpha/2}$ = the bold numbers on the side and top of the table

Ans) $z_{\alpha/2} = 1.96$

P Value Mean Example (σ known)

An insurance company is reviewing its current policy rates. When originally setting the rates they believed that the average claim amount was $1,800. They are concerned that the true mean is actually higher than this, because they could potentially lose a lot of money. They randomly select 40 claims, and calculate a sample mean of $1,950. Assuming that the standard deviation of all claims is $500, and set $\alpha = .05$, test to see if the insurance company should be concerned.

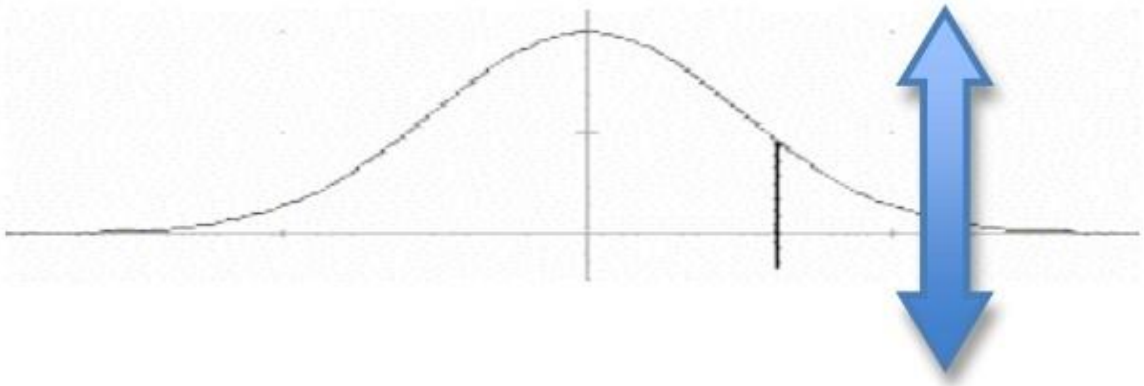
Traditional Method First: Claim: $\mu > \$1800$, so $H_0: \mu = 1800$ & $H_1: \mu > 1800$

Info: $n = 40$, $\bar{x} = 1950$

$\alpha = 0.05$

Critical z-value: 1.645

$\sigma = 500$



$Z_c = 1.645$

Test Statistic: $z = \dfrac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \dfrac{1950 - 1800}{500/\sqrt{40}} = 1.90$, the test statistic falls in the critical region so we ***reject*** H_0.

An insurance company is reviewing its current policy rates. When originally setting the rates they believed that the average claim amount will be maximum Rs180000. They are concerned that the true mean is actually higher than this, because they could potentially lose a lot of money. They randomly select 40 claims, and calculate a sample mean of Rs195000. Assuming that the standard deviation of claims is Rs50000 and set $\alpha = .05$, test to see if the insurance company should be concerned or not.

$Ans)\ z_{cal} = 1.89$

Trying to encourage people to stop driving to campus, the university claims that on average it takes at least 30 minutes to find a parking space on campus. I don't think it takes so long to find a spot. In fact I have a sample of the last five times I drove to campus, and I calculated *x = 20. Assuming that the time* it takes to find a parking spot is normal, and that *σ = 6 minutes, then perform a hypothesis test with level α= 0.10 to see if my claim is correct.*

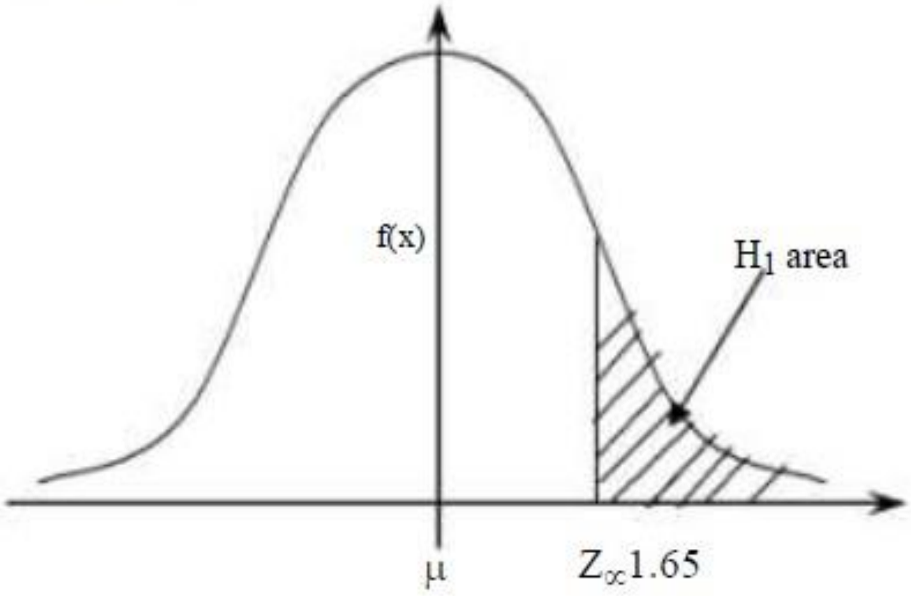
$Ans)\ z_{cal} = -3.726$

Suppose it is known that the mean annual income of workers in a certain community is Rs. 5,000 with a standard deviation of Rs. 1,200. A researcher suspects the "son of the soil" workers have higher than the average income. He draws a random sample of 144 local workers and obtains the sample mean of Rs. 5,500. Can he say that the local workers have significantly higher income than the total population ? (Use $\alpha = 0.05$).

The null and alternative hypotheses are as follows:

H_0: $\mu = 5000$ Have same income

H_1: $\mu > 5000$ Have higher income



This is a one tailed test. Diagram

Given $\bar{x} = 5500$

$n = 144$

$\sigma = 1200$

$\propto = 0.05$ then the test statistic is $Z = \dfrac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \dfrac{5500 - 5000}{1200 / \sqrt{144}} = 5$

From normal distribution tables Z at $\propto = 0.05$ is 1.65. Then, since calculated $Z > 1.65$ then we reject H_0 accept H_1. Meaning the local workers have significantly higher income than the total population.

The Nagari hospital uses large quantities of packaged doses of a particular drug. The individual doses of this drug is 120 cubic centimetres (120 c.c.). The action of the drug is such that the body will harmlessly pass off the excessive doses, and the insufficient doses do not produce the desired medical effect. The hospital, on the basis of past purchases, has estimated the population standard deviation of 3 c.c. It inspects 50 doses of this drug at random from a large shipment and finds the mean of these doses as 119 c.c. The hospital authorities suspect that the dosages in this shipment are too small. Will you agree? (Take $\alpha = 0.10$).

$Ans)\ z_{cal} = -2.357$

A radio shop sells, on an verage, 200 radios per day with a standard deviation of 50 radios After an extensive advertising campaign, the management will compute the average sales for the next 25 days to see whether an improvement has occurred. Assume that the daily sales of radios is normally distributed.

(i) *Write down the null and the alternative hypotheses.*

(ii) *Test the hypothesis at 5% level of significance if $\bar{X} = 216$.*

(iii) *How large must $\bar{X}$ be in order that the null hypothesis is rejected at 5% level of significance.*

$Ans)\ z_{cal} = 1.6$

Solution :-

given that

population mean $(\mu) = 200$

standard deviation $(\sigma) = 50$

sample size $(n) = 25$

sample mean $(\bar{x}) = 216$

(a) Null Hypothesis $H_0: \mu = 200$

Alternative Hypothesis $H_A: \mu > 200$ (Right tailed test).

(b) At $\alpha = 0.05$ level of significance and if sample mean $(\bar{x}) = 216$

test statistic -

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{216 - 200}{50/\sqrt{25}}$$

$$\boxed{Z = 1.6}$$

At $\alpha = 0.05$ the p value can be computed Using R software -

```
> pnorm(1.6, 0, 1, lower.tail =TRUE)
[1] 0.0548
```

A weighing machine without any display was used by an average of 320 persons a day with a standard deviation of 50 persons. When an attractive display was used on the machine, the average for 100 days increased by 15 persons. Can we say that the display did not help much ? Use a level of significance of 0·05.

Hint

In the usual notations, we are given that

$\mu = 320, \quad \sigma = 50,$

$n = 100, \quad \bar{x} = 320 + 15 = 335.$

Null Hypothesis, we set up the null hypothesis

$H_0 : \mu = 320$

Examples:

1. Ten individuals are chosen at random from normal population and their heights are found to be 63, 63, 66, 67, 68, 69, 70, 70, 71, 71 inches. Test if the sample belongs to the population whose mean height is 66 at 5% significant level.

Solution:

$n = 10$, $\mu = 66$ (Population mean)

H_0: $\mu1$ (sample mean) $= \mu$

H_A: $\mu1 != \mu$ (Two Tailed)

- $\mu 1 = 67.8$, $\sigma = 2.86$

$$Z = \frac{\mu_1 - \mu}{\sigma / \sqrt{n}} = 1.99$$

Z= 1.99 > 1.96 (critical value at 5% level of significance)

Hence, H_0 is rejected.

A samples of 50 pieces of certain type of string was tested. The mean breaking strength turned out to be 14.5 kgs. Test whether the sample is from a batch of strings having a mean breaking strength of 15.6 kgs and standard deviation of 2.2 kgs.

Here, Number of sample, $n = 50$

Sample mean $\bar{x} = 14.5$ kgs.

Population mean $\mu = 15.6$ kgs

Standard deviation $\sigma = 2.2$ kgs

Null Hypothesis: $H_o : \mu = 15.6$ kgs i.e. the mean breaking strength of the strings is 15.6 kg.

Alternative Hypothesis: $H_1 : \neq 15.6$ kgs i.e. the mean breaking strength is not equal to 15.6 kg. (two tailed test)

$$\text{We have, } Z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \frac{14.5 - 15.6}{2.2 / \sqrt{50}} = \frac{-1.1}{0.31} = -3.55$$

$$\therefore |z| = |-3.55| = 3.55$$

Level of significance $\alpha = 5\%$

Tabulated value of z at 5% level of significance for two tailed test = 1.96

Since $Z_{cal} > Z_{tab}$, H_o is rejected i.e. H_1 is accepted. Hence we conclude that the sample has not been drawn from the normal population with mean 15.6 kgs and standard deviation 2.2 kgs.

In the past a blending process has produced an average of 5 Kg. of waste material for every batch with standard deviation 5 Kg. From a sample of 100 batches an average of 8 Kg. of waste per batch is obtained. At 5% level of significance is it reasonable to believe that the average has increased?

An insurance agent has claimed that the average age of policyholders who insure through him is less than the average for all agents, which is 30·5 years.

A random sample of 100 policyholders who had insured through him gave the following age distribution :

Age last birthday	*No. of persons*
16—20	12
21—25	22
26—30	20
31—35	30
36—40	16

Calculate the arithmetic mean and standard deviation of this distribution and use these values to test his claim at the 5% level of significance. You are given that Z (1·645) = 0·95.

We need to calculte the sample mean and sample standard deviation

LL	UL	f	xi	$di = \frac{xi - A}{h}$	fidi	$fi(di^2)$
16	20	12	18	-2	-24	48
21	25	22	23	-1	-22	22
26	30	20	28	0	0	0
31	35	30	33	1	30	30
36	40	16	38	2	32	64
$\sum$		100			16	164

A=	28	
h=	5	
N=	100	

$$\bar{x} = A + \frac{h \times \sum fd}{N}$$

where N= $\sum f$

$$s = h \times \sqrt{\frac{\sum fd^2}{N} - \left(\frac{\sum fd}{N}\right)^2}$$

$\bar{x}$ = 28 + $\dfrac{5 \times 16}{100}$ $\qquad \bar{x} = A + \dfrac{h \times \sum fd}{N}$

$\bar{x}$ = 28.8

$\sum fd^2/N$ = 1.64

$\sum fd/N$ = 0.16

$(\sum fd/N)^2$ = 0.0256

s= 6.35295 $\qquad s = h \times \sqrt{\dfrac{\sum fd^2}{N} - \left(\dfrac{\sum fd}{N}\right)^2}$

Thus the sample standard deviation is 6.35 which can be taken as population standard deviation σ as the sample is large

$\bar{x}$ = 28.8

σ= 6.353

as n>=30 , large sample we can use z test for single mean

Standard Error =	0.6353
Test statistic Z=	$\frac{-1.7}{0.6353}$
Test statistic Z =	-2.676
|z|=	2.68
Zcritical =	1.645

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

	CRITICAL VALUES OF Z			
	LEVEL OF SIGNIFICANCE (α)			
	1%	2%	5%	10%
Two-tailed test	$\lvert z_{\alpha/2}\rvert$=2.58	$\lvert z_{\alpha/2}\rvert$=2.33	$\lvert z_{\alpha/2}\rvert$=1.96	$\lvert z_{\alpha/2}\rvert$=1.645
Right-tailed test	z_α=2.33	z_α=2.05	z_α=1.645	z_α=1.28
Left-tailed test	$-z_\alpha$=-2.33	$-z_\alpha$=-2.05	$-z_\alpha$=-1.645	$-z_\alpha$=-1.28

As zcalc > zcritical, we reject the Null Hypothesis at 0.05 LOS or 95% confidence

Thus we reject the null Hypothesis that "Average age of policy holders who insure through the agent is μ = 30.5" at 5 % LOS

Conclusion : Thus the insurance agent claim that the average age of policy holders who insure through him is > 30.5 is true at 5 % LOS

An insurance agent has claimed that the average age of policy-holders who insure through him is less than the average for all agents, which is 32 years. A random sample of 100 policy-holders who had insured through him gave an average age of 30 years. Assuming a standard error of 5 years, do you think that his claim is justifiable? Use α at 5% level of significance to test the claim.

The null hypothesis, $H_0 : \mu = 32$ years and the alternative hypothesis $H_1 : \mu < 32$ years.

$$Z = \frac{X - \mu}{\sigma_{\bar{x}}}$$

$$= \frac{30 - 32}{\boxed{5}}$$

$$= -0.4$$

We are interested in testing whether or not the insurance agent's claim of average age of policy-holders who insure through him is justified. Thus, it is a left-tail test as the alternative hypothesis $\mu < 32$ years.

At $\alpha = 0.05$ level of significance, the critical value of Z is -1.64 for a one-tail test. The computed value of $Z = -0.4$ falls in the acceptance region, as shown in Fig. 14.2. Thus, we accept the null hypothesis and conclude that the claim made by the insurance agent is not justifiable.

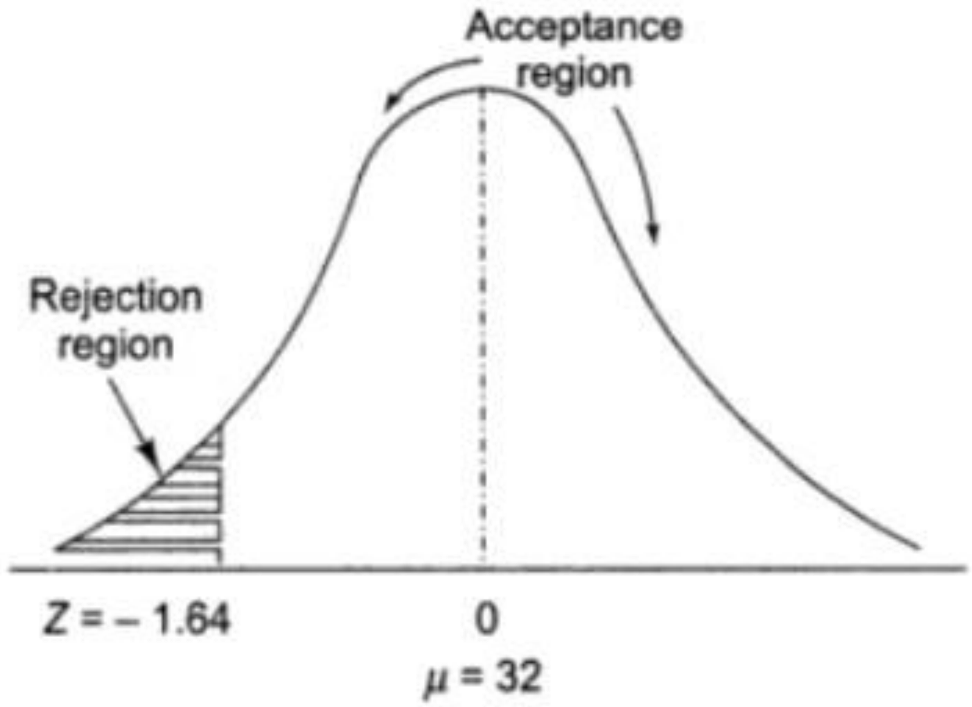


Fig. 14.2 Test of Hypothesis
$H_0 : \mu = 32$ years

A random sample of 400 items is found to have a mean of 82 and standard deviation of 18. Find 95% confidence limits for the mean of the population from which the sample is drawn.

we are given:

$n = 400, \bar{x} = 82$ and

$S = 18.$

95% confidence limits for the population mean μ are given by:

$\bar{x} \pm 1.96 \times \sigma/\sqrt{n} \doteq \bar{x}$

Given that the standard deviation of household expenditure from a pilot survey is Rs. 7·2, what minimised sample should be taken to ascertain the mean level of expenditure so that we can be 95% confident that the population mean expenditure lies within Rs. 2·00 either way of the sample mean expenditure ?

A sample of size 400 was drawn and the sample mean was found to be 99. Test whether this sample could have come from a normal distribution with mean 100 and variance 64 at 5% LOS

Hypothesis Testing: Z test for difference of mean

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sigma\sqrt{(1/n_1 + 1/n_2)}}$$

The means of two single large samples of 1000 and 2000 members are 67·5 inches and 68·0 inches respectively. Can the samples be regarded as drawn from the same population of standard deviation 2·5 inches? (Test at 5% level of significance).

α %	15 %	10 %	5 %	4 %	1%	0.5 %	0.2 %
α	0.15	0.1	0.05	0.04	0.01	0.005	0.002
$-Z_{\alpha/2}$ and $+Z_{\alpha/2}$ for T.T.T.	− 1.44 1.44	− 1.645 1.645	− 1.96 1.96	− 2.06 2.06	− 2.58 2.58	− 2.81 2.81	− 3.08 3.08
$-Z_{\alpha}$ for L.O.T.T.	− 1.04	− 1.28	− 1.645	− 2.6	− 2.33	− 2.58	− 2.88
Z_{α} for R.0.T.T.	1.04	1.28	1.645	2.6	2.33	2.58	2.88

Solution:

Given:

n1=	1000	
n2=	2000	
x̄1 =	67.5	inches
x̄2 =	68	inches

H0: μ 1=μ2 samples are drawn from the same population with standard deviation σ =2.5 inches

Ha: μ 1≠ μ2

σ = 2.5 inches

x̄1 - x̄2 = -0.5

σ *sqrt(1/n1+1/n2) = 0.0968

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sigma\sqrt{(1/n_1 + 1/n_2)}}$$

z = -5.164

|z|= 5.164

Conclusion: Since I Z I > 3, the value is highly significant and, we reject the null hypothesis and conclude that samples are cenainly not from the same population with standard deviation 2·5.

Hypothesis Testing: Z test for difference of mean

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

The mean height of 50 male students who showed interest in college athletics was 68.2 inches with a standard deviation of 2.5 inches, while 50 male students who showed no interest in athletics had a mean height of 67.5 inches with a S.D. of 2.8 inches. Test the hypothesis that male students who showed interest in athletics are taller than other male students.

Solution:

Given:

$n1=$	50	
$n2=$	50	
$\bar{x}1 =$	68.2	inches
$\bar{x}2 =$	67.5	inches
$\sigma 1=$	2.5	inches
$\sigma 2=$	2.8	inches

H0: $\mu\ 1=\mu 2$ there is not significant differene between mean heights of students who participate and and do not participate in the

Ha: $\mu\ 1 > \mu 2$, male students who participate in college athletics are taller than other students

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

$\bar{x}1 - \bar{x}2 =$	0.7
$\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$	0.53085
$z =$	1.31864
$\lvert z \rvert =$	1.31864
Zcritical @ 5% LOS	1.645

CRITICAL VALUES OF Z

	LEVEL OF SIGNIFICANCE (α)			
	1%	2%	5%	10%
Two-tailed test	$\lvert z_{\alpha/2} \rvert=2.58$	$\lvert z_{\alpha/2} \rvert=2.33$	$\lvert z_{\alpha/2} \rvert=1.96$	$\lvert z_{\alpha/2} \rvert=1.645$
Right-tailed test	$z_{\alpha}=2.33$	$z_{\alpha}=2.05$	$z_{\alpha}=1.645$	$z_{\alpha}=1.28$
Left-tailed test	$-z_{\alpha}=-2.33$	$-z_{\alpha}=-2.05$	$-z_{\alpha}=-1.645$	$-z_{\alpha}=-1.28$

Conclusion: As **zcal<zcritical** we fail to reject null hypothesis. We cannot conclude that male students who participate in college athletics are taller than other students. i.e. average height of male students who participate in college athletics is same as other male students.

Hypothesis Testing: Z test for difference of mean (Homework)

In order to make a survey of the buying habits, two markets A and B are chosen at two different parts of a city. 400 women shoppers are chosen at random in market A. Their average daily expenditure on food is found to be Rs. 250 with a standard deviation of Rs. 40. The figures are Rs. 220 and Rs. 55 respectively in the market B where also 400 women shoppers are chosen at random. Test at 1% level of significance, whether the average daily food expenditures of the two shoppers are equal.

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$