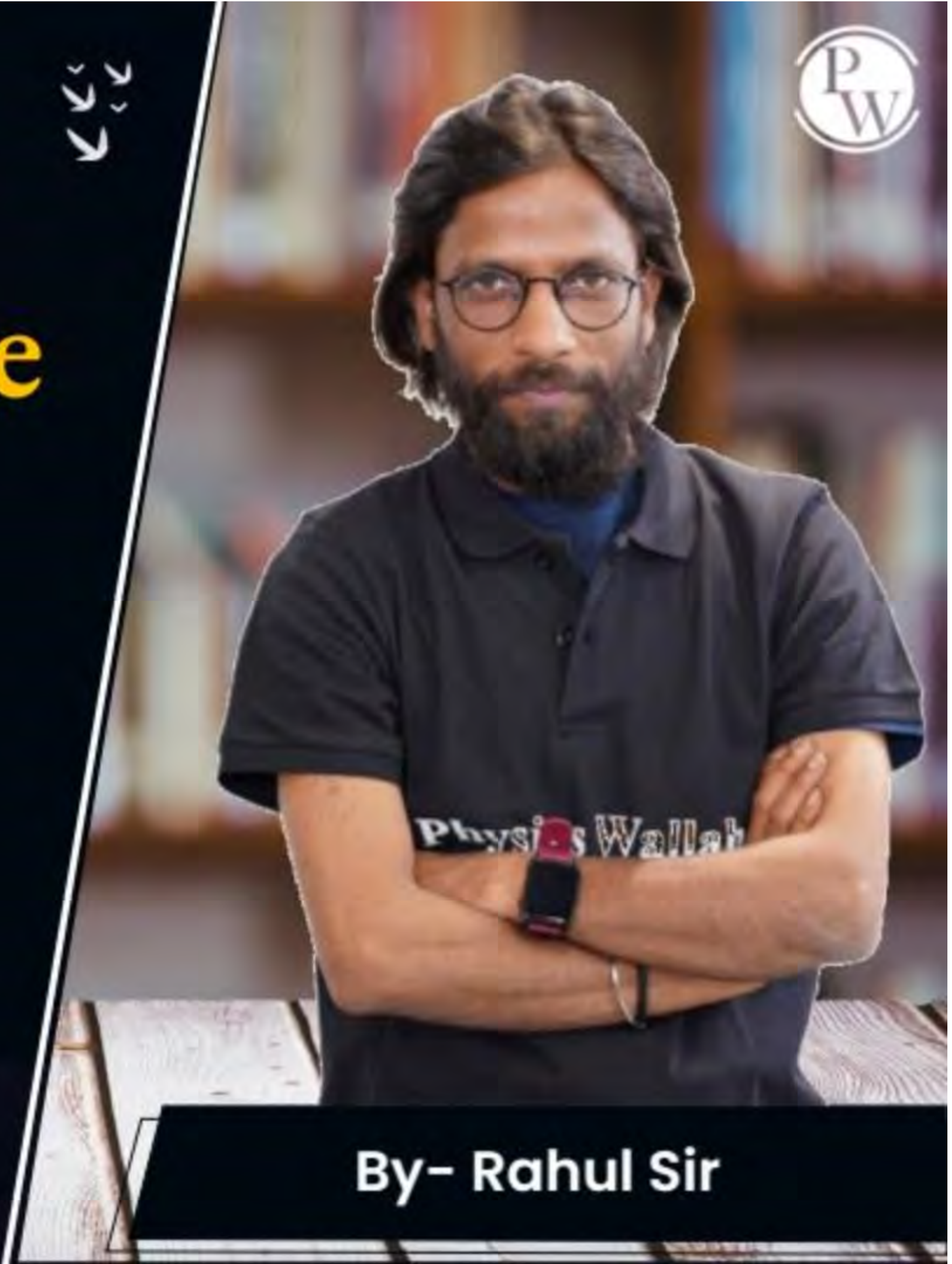
Data Science and Artificial Intelligence

Probability and Statistics

Testing of Hypothesis

Lecture No.- 02



By- Rahul Sir

Topics to be Covered

Topic Introduction to t-test

Introduction to t-TEST :-

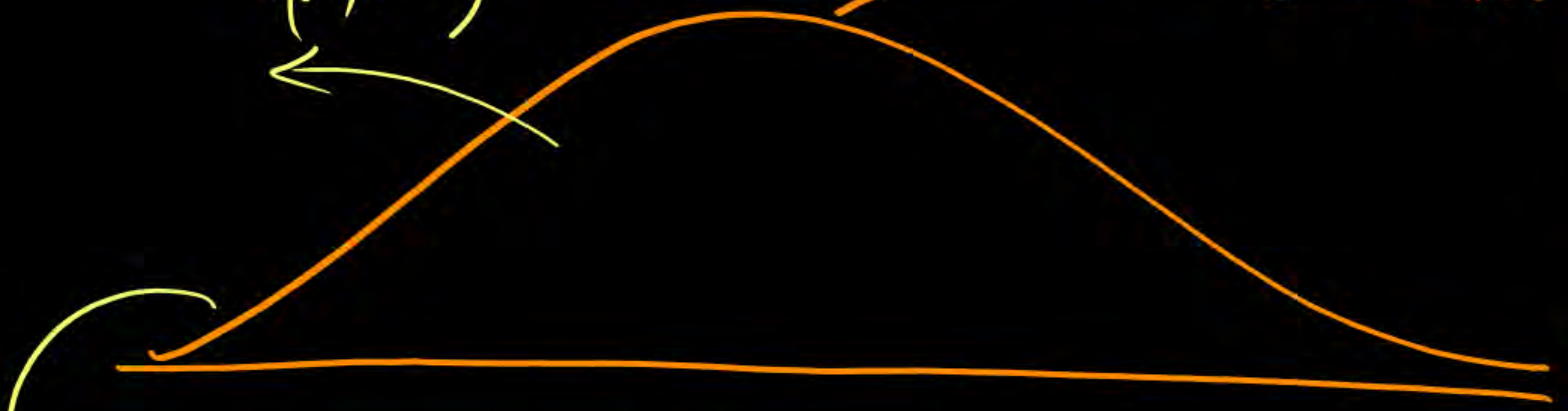
– Goodness of fit / degree of Freedom 'n'

$N(\mu, \sigma^2)$

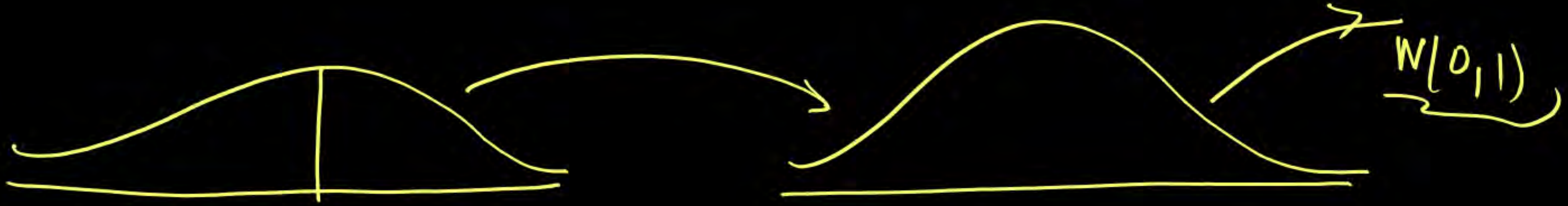
t-distribution $n < 30$

Using t-TEST

T-test allowed The sample size $n < 30$



almost Look like – Gaussian random var



T-test $\Rightarrow$ $T = \dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} = \dfrac{\bar{X}-\mu}{\sigma/\sqrt{\text{sample size}}}$ → SAMPLE $= \dfrac{1}{n-1}\sum(x_i-\bar{x})^2$

$\bar{X}$ = SAMPLE mean

μ = Population mean

σ = Sample variance

$$\sigma = \frac{1}{(n-1)}\sum_{i=1}^{n}(x_i-\bar{x})^2$$

If Null Hypothesis is Accept

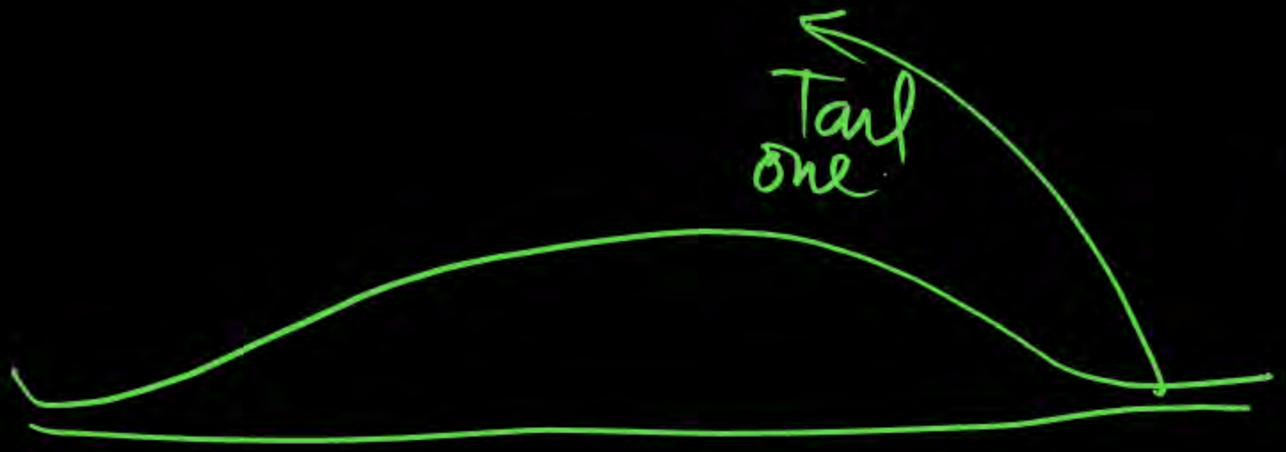
Test statistic $\Rightarrow$ one Tailed Test

T value < (degree of Freedom $_{0.05}$)

T value > (degree of freedom)$_{0.05}$

→ Null Hypothesis – rejected

Using Alternate Hypothesis → Accept



t-distribution and t-test

Q1. A random sample of size 16 has 53 as mean. The sum of squares of the deviation from mean is 135. Can this sample be regarded as taken from the population having 56 as mean? Obtain 95% and 99% confidence limits of the mean of the population.

$n = 16$ (sample size)

Population mean $\mu = 56$

sample mean $\bar{x} = 53$

$\sum (x_i - \bar{x})^2 = 135$

Using T-test

$$T = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{53 - 56}{}$$

$$\sigma^2 = \frac{1}{16-1} \times 135 = \frac{135}{15} =$$

$$\sigma = \sqrt{\frac{135}{15}} = \bigcirc$$

Slide 3

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

$$\sigma^2 = \frac{1}{(n-1)} \sum (x_i - \bar{x})^2$$

$$= \frac{1}{(16-1)} \times 135$$

$$\frac{-3}{\frac{3}{\sqrt{16}}}$$

$$= -\frac{\cancel{3}}{\frac{\cancel{3}}{4}}$$

$$= -\underline{\underline{4}}$$

$$T = \frac{53-56}{3/\sqrt{16}} \Rightarrow -4 \checkmark$$

T–

$\rightarrow$ Test value always positive

σ $\underline{\underline{S.D}}$ $= \sqrt{\frac{135}{15}}$

$\checkmark$ $T = |-4| = 4$

$\checkmark$ degree of freedom $= 16 - 1 = 15$

Rejected Hypothesis — Null Hypothesis reject

If 95% confidence Interval

$$\text{sample mean} = \bar{X} + \frac{\sigma}{\sqrt{n}} t_{0.05}$$

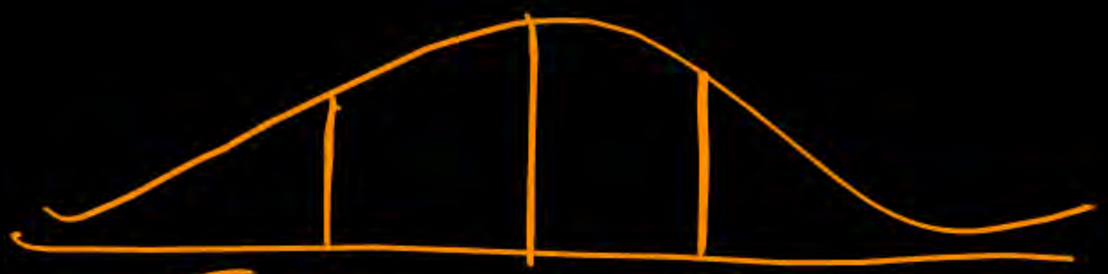
Sample mean
S.D

$$\text{sample mean} = \bar{X} - \frac{\sigma}{\sqrt{n}} t_{0.05}$$

95%

If 99% confidence Interval

$$\text{Sample mean} = \bar{X} \pm \frac{\sigma}{\sqrt{n}} t_{0.01}$$



$$\text{Sample mean} = 53 \pm \frac{3}{\sqrt{16}} \times t_{0.05} = 51. \quad , 54.$$

$$= 53 \pm \frac{3}{\sqrt{16}} \times t_{0.01} = 50. \quad , 55.$$

t-distribution and t-test

$\boxed{T = 2.123}$ → T-test

Q2. The lifetime of electric bulbs for a random sample of 10 from a large consignment gave the following data:

Item	1	2	3	4	5	6	7	8	9	10
Life in 000 hrs	4.2	4.6	3.9	4.1	5.2	3.8	3.9	4.3	4.4	5.6

Can we accept the hypothesis that the average lifetime of bulb is 4000 hrs?

$\boxed{\text{Accepted The Hypothesis}}$ → $\bar{x} \pm \frac{\sigma}{\sqrt{n}} t_{0.05}$ — 95%

$\bar{x} \pm \frac{\sigma}{\sqrt{n}} t_{0.01}$ — 99%

t-distribution and t-test

Q3. A sample of 20 items has mean 42 units and S.D. 5 units. Test the hypothesis that it is a random sample from a normal population with mean 45 units.

$n = 20$

$\bar{x}$ mean $= 42$ units

σ S.D $= 5$ units

Population mean $= 45$ units

H_0 Null Hypothesis $\mu = 45$

H_1 Alternate $\mu \neq 45$

one tail Test

Two Tail test

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

$$\sigma^2 = \frac{1}{n-1} \sum (x_i - x)^2$$

$$T = \frac{42-45}{5/\sqrt{19}} = -2.615$$

$$\sigma^2 = \frac{\sum (x_i - x)^2}{(n-1)}$$

$$|T| = |-2.615| = 2.615$$

Degree of Freedom $= (20-1) = 19$

$$t_{0.05} = \underline{2.093}$$

<u>Reject The Hypothesis</u>

t-distribution and t-test

Q4. The 9 items of a sample have the following values

45, 47, 50, 52, 48, 47, 49, 53, 51.

Does the mean of these values differ significantly from the assumed mean 47.5?

$$\bar{X} = \frac{45+47+50+52+48+47+49+53+51}{9} = 49.11$$

$\mu = 47.5$

X	45	47	50	52	48	47	49	53	51
$\sum(X-\bar{X})^2$									

$\bar{X} = 49.11$

$S.D = \frac{1}{(n-1)}\sum(X-\bar{X})^2$

σ^2

$\boxed{\sigma = (2.622)}$

T test

$$T = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} = \frac{49.11-47.5}{2.622/\sqrt{9}}$$

$T = 1.83$

$\boxed{|T| = 1.83}$

Hypothesis Accepted

$$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$$

Sample mean

Population

$$\sigma^2 = \frac{1}{(n-1)} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

$t_{0.05,15} = 1.753$

alternate

Null

t Table

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$	$t_{.999}$	$t_{.9995}$
one-tail	0.50	0.25	0.20	0.15	0.10	0.05	0.025	0.01	0.005	0.001	0.0005
two-tails	1.00	0.50	0.40	0.30	0.20	0.10	0.05	0.02	0.01	0.002	0.001
df											
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	318.31	636.62
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.000	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.000	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.000	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.000	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.000	0.691	0.866	1.074	1.341	$\boxed{1.753}$	2.131	2.602	2.947	3.733	4.073
16	0.000	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.686	4.015

17	0.000	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.000	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.000	0.688	0.861	1.066	1.328	1.729	$\boxed{2.093}$	2.539	2.861	3.579	3.883
20	0.000	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	0.000	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.000	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	0.000	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	0.000	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	0.000	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	0.000	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	0.000	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	0.000	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	0.000	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	0.000	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	0.000	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	0.000	0.679	0.848	1.045	1.296	1.671	2.000	2.390	2.660	3.232	3.460
80	0.000	0.678	0.846	1.043	1.292	1.664	1.990	2.374	2.639	3.195	3.416
100	0.000	0.677	0.845	1.042	1.290	1.660	1.984	2.364	2.626	3.174	3.390
1000	0.000	0.675	0.842	1.037	1.282	1.646	1.962	2.330	2.581	3.098	3.300
z	0.000	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576	3.090	3.291
	0%	50%	60%	70%	80%	90%	95%	98%	99%	99.8%	99.9%
	Confidence Level										

THANK - YOU