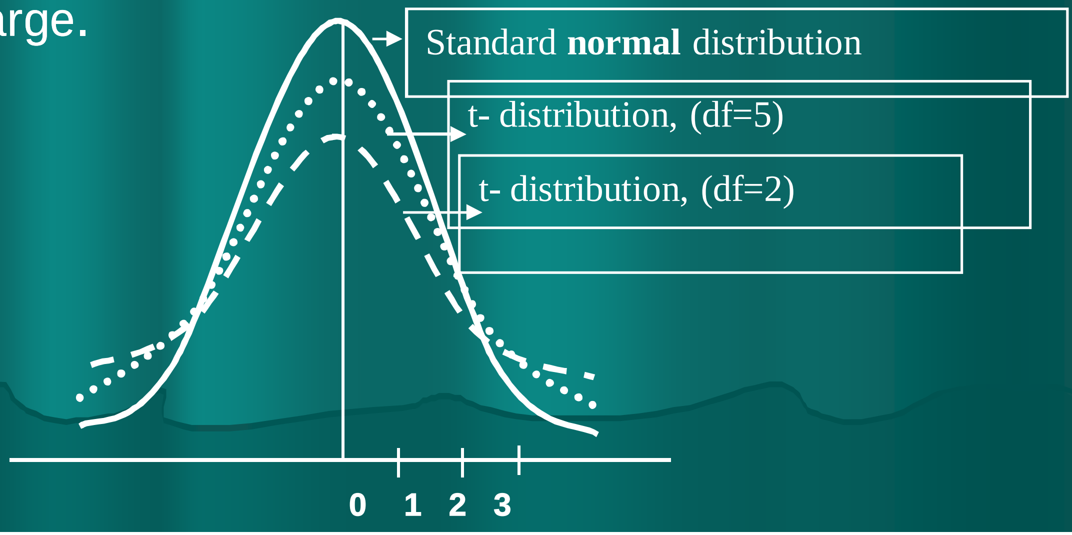
t-Distribution

- In biological research the population *variance is usually unknown and an unbiased estimate s2*, obtained from the sample data, has to be used in place of σ^2.
- The properties of t- distribution are:
- 1-It is symmetrical about the mean.
- 2-It is really family distributions, which depend on the degrees of freedom (df), of samples.
- 3-It is not normal distribution and it approach to normal as (n) approaches infinity.
- 4-t curve is somewhat flatter than that for standardized normal distribution.
- 5-It has a mean zero and variance more than one but it approaches one as the sample size become large.



- Confidence Interval for Mean μ, when σ^2 is *Unknown*, and n≤30 (C.I):

  A (1-a) 100% C.I. for is given by

$$\bar{x} \pm t_{(1-a/2),v}\frac{s}{\sqrt{n}} \quad , \quad v = df$$

- *From percentiles of the t-distribution table construct this value,*

$$t_{(1-a/2),v}$$

- Example: A sample of 16 ten-year-old girls given a mean weight of 71.5 and standard deviation of 12 pounds, respectively. Assuming normality. Find 95 percent confidence interval for, (μ is true population mean of the weight girls).
- A 95% C.I. for μ is,

$$71.5 \pm 2.1315\frac{12}{\sqrt{16}}$$

$$(65.106 \quad , \quad 77.895)$$

- *Meaning that based on sample statistics we are 95% confident that the true population mean of the weight girls lies within the range of 65.106 to 77.895 pound.*

- Note: *When the sample n > 30, our faith in (s) as an approximation of (s) is usually substantial, and we may feel justified in using normal distribution theory to construct confidence interval.*
- Confidence Interval for the Difference between Two Population Means μ_1 and μ_2; When $\sigma_1 = \sigma_2$ but are unknown, (C.I);
-
- A (1- a) 100% confidence interval for μ_1 - μ_2 is given by,
- $$(\bar{x}_1 - \bar{x}_2) \pm t_{(1-\alpha/2),v}\, Sp\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} \quad , v = n_1 + n_2 - 2$$
- $$Sp = \sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}}$$ (Pooled estimate of standard deviation)

- Example: The following data represent the fasting blood glucose of samples for normal males and females. Assume that the fasting blood glucose of variances are equals ($\sigma_1^2 = \sigma_2^2$), but unknown. Construct 95 percent confidence interval for the difference between the two true population means of the fasting blood glucose.
- Blood Glucose (mg/dl)
- Male: 97 100 115 98 110 105 115
- Female: 103 94 110 90 100

$$\bar{x}_1 = 105.714 \quad , \quad s_1 = 7.739$$

$$\bar{x}_2 = 99.4 \quad , \quad s_2 = 7.797$$

$$Sp = \sqrt{\frac{(7-1)7.739^2 + (5-1)7.797^2}{7+5-2}} = 7.762$$

95% C.I. for μ_1 - μ_2 is, $(\bar{x}_1 - \bar{x}_2) \pm t_{(1-a/2),v}\, Sp\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \quad , v = n_1 + n_2 - 2$

$$(105.714 - 99.4) \pm (2.2281)(7.762)\sqrt{\frac{1}{7} + \frac{1}{5}}$$

$$(-3.813 \quad , \quad 16.441)$$

- *Meaning that based on sample statistics we are 95% confident that the true difference between two true population means of the two groups of fasting blood glucose lies within the range of –3.813 to 16.441.*

Confidence Interval for the Difference between Two Population Means μ_1 and μ_2;when $\sigma_1 \neq \sigma_2$ but they are unknown, (C.I);

An approximate (1 -α) 100% confidence interval for $\mu_1 - \mu_2$ is given by,

$$\left(\bar{x}_1 - \bar{x}_2\right) \pm t_{(1-a/2),v'} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$or\ P\left[\left(\bar{x}_1 - \bar{x}_2\right) - t_{(1-\alpha/2),v'} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \leq \mu_1 - \mu_2 \leq \left(\bar{x}_1 - \bar{x}_2\right) + t_{(1-\alpha/2),v'} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}\right] = 1-\alpha$$

$$df' = v' = \frac{\left(s_1^2/n_1 + s_2^2/n_2\right)^2}{\left[\left(s_1^2/n_1\right)^2/(n_1-1)\right] + \left[\left(s_2^2/n_2\right)^2/(n_2-1)\right]}$$

- Example: Total serum complement activity (CH_{50}) was assayed in 20 apparently healthy subjects and 10 subjects with disease. The following result were obtained:

Subjects	n	$\bar{x}$	s
With disease	10	62.6	33.8
Normal	20	47.2	10.1

- The investigator had reason to believe that the sampled populations are approximately normally distributed, but they were unwilling to assume that the two unknown population variances are equal. Find 95 percent confidence interval for $\mu_1 - \mu_2$.
- An approximate 95% C.I. for $\mu_1 - \mu_2$.

$$v' = \frac{\left(\frac{33.8^2}{10} + \frac{10.1^2}{20}\right)^2}{\frac{\left(\frac{33.8^2}{10}\right)^2}{10-1} + \frac{\left(\frac{10.1^2}{20}\right)^2}{20-1}} = 9.8$$

$$15.4 \pm 2.228\sqrt{\frac{33.8^2}{10} + \frac{10.1^2}{20}}$$

$$15.4 \pm 24.4$$

$$(-9.0 \quad , \quad 39.8)$$

- *Meaning that based on sample statistics we are 95% confident that the true difference between two true population means of the two groups lies within the range of – 9.0 to 39.8.*

Percentiles of the Student's t distribution

	Probability that a larger value of t would be observed = 1 minus the values shown							
df	0.6	0.7	0.8	0.9	0.95	0.975	0.99	0.995
1	0.325	0.727	1.367	3.078	6.314	12.706	31.821	63.657
2	0.289	0.617	1.061	1.886	2.92	4.303	6.965	9.925
3	0.277	0.584	0.978	1.638	2.353	3.182	4.541	5.841
4	0.271	0.569	0.941	1.533	2.132	2.776	3.747	4.604
5	0.267	0.559	0.92	1.476	2.015	2.571	3.365	4.032
6	0.265	0.553	0.906	1.44	1.943	2.447	3.143	3.707
7	0.263	0.549	0.896	1.415	1.895	2.365	2.998	3.499
8	0.262	0.546	0.889	1.397	1.86	2.306	2.896	3.355
9	0.261	0.543	0.883	1.383	1.833	2.262	2.821	3.25
10	0.26	0.542	0.879	1.372	1.812	2.228	2.764	3.169
11	0.26	0.54	0.876	1.363	1.796	2.201	2.718	3.106
12	0.259	0.539	0.873	1.356	1.782	2.179	2.681	3.055

13	0.259	0.538	0.87	1.35	1.771	2.16	2.65	3.012
14	0.258	0.537	0.868	1.345	1.761	2.145	2.624	2.977
15	0.258	0.536	0.866	1.341	1.753	2.131	2.602	2.947
16	0.258	0.535	0.865	1.337	1.746	2.12	2.583	2.921
17	0.257	0.534	0.863	1.333	1.74	2.11	2.567	2.898
18	0.257	0.534	0.862	1.33	1.734	2.101	2.552	2.878
19	0.257	0.533	0.861	1.328	1.729	2.093	2.539	2.861
20	0.257	0.533	0.86	1.325	1.725	2.086	2.528	2.845
21	0.257	0.532	0.859	1.323	1.721	2.08	2.518	2.831
22	0.256	0.532	0.858	1.321	1.717	2.074	2.508	2.819
23	0.256	0.532	0.858	1.319	1.714	2.069	2.5	2.807
24	0.256	0.531	0.857	1.316	1.708	2.06	2.485	2.787
25	0.256	0.531	0.856	1.316	1.708	2.06	2.485	2.787
26	0.256	0.531	0.856	1.315	1.706	2.056	2.479	2.779
27	0.256	0.531	0.855	1.314	1.703	2.052	2.473	2.771
28	0.256	0.53	0.855	1.313	1.701	2.048	2.467	2.763
29	0.256	0.53	0.854	1.31	1.697	2.042	2.457	2.75
30	0.256	0.53	0.854	1.31	1.697	2.042	2.457	2.75
40	0.255	0.529	0.851	1.303	1.684	2.021	2.423	2.704
60	0.254	0.527	0.848	1.296	1.671	2	2.39	2.66
120	0.254	0.526	0.845	1.289	1.658	1.98	2.358	2.617
∞	0.253	0.524	0.842	1.282	1.645	1.96	2.326	2.576