Statistics Part IV
Confidence Limits and Hypothesis Testing

Joseph J. Nahas

University of Notre Dame

Statistic Outline (cont.)

3. Graphical Display of Data
   - A. Histogram
   - B. Box Plot
   - C. Normal Probability Plot
   - D. Scatter Plot
   - E. MatLab Plotting
4. Confidence Limits and Hypothesis Testing
   - A. Student's t Distribution
     - i. Who is "Student"
     - ii. Definitions
   - B. Confidence Limits for the Mean
   - C. Equivalence of two Means
   - D. Equivalence of two Variances

Student's t Distribution

- Suppose a random sample of size n is drawn from a normal N(μ,σ) population.
- If $\bar{x}$ is the estimate of the mean from the sample, and s is the sample standard deviation, then

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

  has the t distribution with n-1 degrees of freedom.
- There is a different t distribution for each sample size n as specified in the degrees of freedom.
- WE ARE NOT GOING TO PROVE THIS!

Who was Student?

- The t-distribution was discovered by William S. Gossett, a statistician employed by the Guinness Brewing Company.
  - He was trying to determine how accurate the data from his small samples were.
- Guinness previously had problems with proprietary information being published so it required Gossett not to publish his discoveries under his own name.
  - Guinness did not want its competitors to know that it was using statistics to improve its beer.
- He published the t-distribution under the pen name "Student" in 1908.
- The distribution is usually referred to as

  Student's t Distribution

Student's t distribution

- The probability density function for the t distribution is:

$$f(x) = \frac{\left(1+\frac{x^2}{\nu}\right)^{\frac{-(\nu+1)}{2}}}{B(0.5,0.5\nu)\sqrt{\nu}}$$

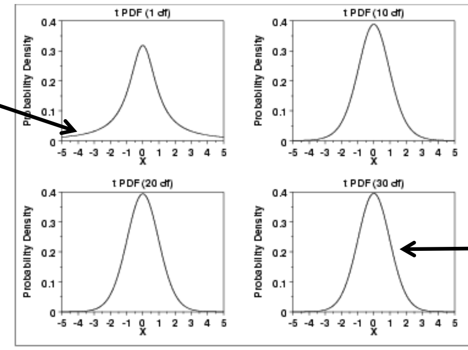
where B is the Beta function and ν is a positive integer shape parameter.

The Beta function is:

$$B(\alpha,\beta) = \int_0^1 t^{\alpha-1}(1-t)^{\beta-1}dt$$

- The t-distribution is equal to the Cauchy distribution for ν = 1.
- The t-distribution approaches the normal distribution for large ν.

t Probability Density Function



Confidence Limits for the Mean

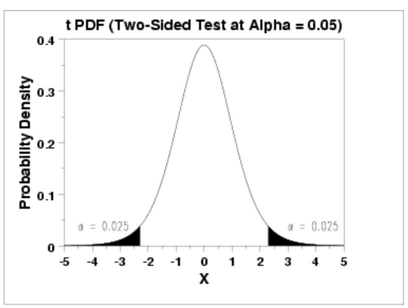
- By definition

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

- So

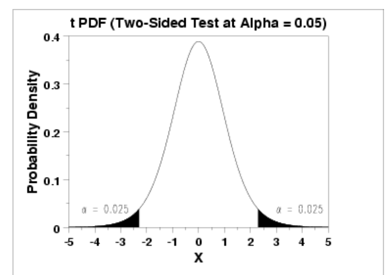
$$\mu = \bar{x} - t_{\alpha,\nu}\frac{s}{\sqrt{n}}$$

- The probably value of μ is distributed around $\bar{x}$.



NIST ESH 1.3.5.2

t-distribution Table Instructions



Given a specified value for α :

1. For a two-sided test, find the column corresponding to 1-α/2 and reject the null hypothesis if the absolute value of the test statistic is greater than the value of $t_{1-\alpha/2,\nu}$ in the table below.
2. For an upper, one-sided test, find the column corresponding to 1-α and reject the null hypothesis if the test statistic is greater than the table value.
3. For a lower, one-sided test, find the column corresponding to 1-α and reject the null hypothesis if the test statistic is less than the negative of the table value.

NIST ESH 1.3.6.7.2

t-distribution table

Probability less than the critical value ($t_{1-\alpha,\nu}$)						
ν	0.90	0.95	0.975	0.99	0.995	0.999
1.	3.078	6.314	12.706	31.821	63.657	318.313
2.	1.886	2.920	4.303	6.965	9.925	22.327
3.	1.638	2.353	3.182	4.541	5.841	10.215
4.	1.533	2.132	2.776	3.747	4.604	7.173
5.	1.476	2.015	2.571	3.365	4.032	5.893
6.	1.440	1.943	2.447	3.143	3.707	5.208
7.	1.415	1.895	2.365	2.998	3.499	4.782
8.	1.397	1.860	2.306	2.896	3.355	4.499
9.	1.383	1.833	2.262	2.821	3.250	4.296
10.	1.372	1.812	2.228	2.764	3.169	4.143
11.	1.363	1.796	2.201	2.718	3.106	4.024
12.	1.356	1.782	2.179	2.681	3.055	3.929
13.	1.350	1.771	2.160	2.650	3.012	3.852
14.	1.345	1.761	2.145	2.624	2.977	3.787
15.	1.341	1.753	2.131	2.602	2.947	3.733
16.	1.337	1.746	2.120	2.583	2.921	3.686
17.	1.333	1.740	2.110	2.567	2.898	3.646
18.	1.330	1.734	2.101	2.552	2.878	3.610
19.	1.328	1.729	2.093	2.539	2.861	3.579
20.	1.325	1.725	2.086	2.528	2.845	3.552
21.	1.323	1.721	2.080	2.518	2.831	3.527
22.	1.321	1.717	2.074	2.508	2.819	3.505
23.	1.319	1.714	2.069	2.500	2.807	3.485
24.	1.318	1.711	2.064	2.492	2.797	3.467
25.	1.316	1.708	2.060	2.485	2.787	3.450
26.	1.315	1.706	2.056	2.479	2.779	3.435
27.	1.314	1.703	2.052	2.473	2.771	3.421
28.	1.313	1.701	2.048	2.467	2.763	3.408
29.	1.311	1.699	2.045	2.462	2.756	3.396
30.	1.310	1.697	2.042	2.457	2.750	3.385

Probability less than the critical value ($t_{1-\alpha,\nu}$)						
ν	0.90	0.95	0.975	0.99	0.995	0.999
31.	1.309	1.696	2.040	2.453	2.744	3.375
32.	1.309	1.694	2.037	2.449	2.738	3.365
33.	1.308	1.692	2.035	2.445	2.733	3.356
34.	1.307	1.691	2.032	2.441	2.728	3.348
35.	1.306	1.690	2.030	2.438	2.724	3.340
36.	1.306	1.688	2.028	2.434	2.719	3.333
37.	1.305	1.687	2.026	2.431	2.715	3.326
38.	1.304	1.686	2.024	2.429	2.712	3.319
39.	1.304	1.685	2.023	2.426	2.708	3.313
40.	1.303	1.684	2.021	2.423	2.704	3.307
41.	1.303	1.683	2.020	2.421	2.701	3.301
42.	1.302	1.682	2.018	2.418	2.698	3.296
43.	1.302	1.681	2.017	2.416	2.695	3.291
44.	1.301	1.680	2.015	2.414	2.692	3.286
45.	1.301	1.679	2.014	2.412	2.690	3.281
46.	1.300	1.679	2.013	2.410	2.687	3.277
47.	1.300	1.678	2.012	2.408	2.685	3.273
48.	1.299	1.677	2.011	2.407	2.682	3.269
49.	1.299	1.677	2.010	2.405	2.680	3.265
50.	1.299	1.676	2.009	2.403	2.678	3.261
51.	1.298	1.675	2.008	2.402	2.676	3.258
52.	1.298	1.675	2.007	2.400	2.674	3.255
53.	1.298	1.674	2.006	2.399	2.672	3.251
54.	1.297	1.674	2.005	2.397	2.670	3.248
55.	1.297	1.673	2.004	2.396	2.668	3.245
56.	1.297	1.673	2.003	2.395	2.667	3.242
57.	1.297	1.672	2.002	2.394	2.665	3.239
58.	1.296	1.672	2.002	2.392	2.663	3.237
59.	1.296	1.671	2.001	2.391	2.662	3.234
60.	1.296	1.671	2.000	2.390	2.660	3.232

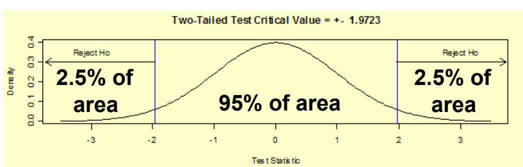
t-distribution table

Probability less than the critical value ($t_{1-\alpha,\nu}$)						
ν	0.90	0.95	0.975	0.99	0.995	0.999
61.	1.296	1.670	2.000	2.389	2.659	3.229
62.	1.295	1.670	1.999	2.388	2.657	3.227
63.	1.295	1.669	1.998	2.387	2.656	3.225
64.	1.295	1.669	1.998	2.386	2.655	3.223
65.	1.295	1.669	1.997	2.385	2.654	3.220
66.	1.295	1.668	1.997	2.384	2.652	3.218
67.	1.294	1.668	1.996	2.383	2.651	3.216
68.	1.294	1.668	1.995	2.382	2.650	3.214
69.	1.294	1.667	1.995	2.382	2.649	3.213
70.	1.294	1.667	1.994	2.381	2.648	3.211
71.	1.294	1.667	1.994	2.380	2.647	3.209
72.	1.293	1.666	1.993	2.379	2.646	3.207
73.	1.293	1.666	1.993	2.379	2.645	3.206
74.	1.293	1.666	1.993	2.378	2.644	3.204
75.	1.293	1.665	1.992	2.377	2.643	3.202
76.	1.293	1.665	1.992	2.376	2.642	3.201
77.	1.293	1.665	1.991	2.376	2.641	3.199
78.	1.292	1.665	1.991	2.375	2.640	3.198
79.	1.292	1.664	1.990	2.374	2.640	3.197
80.	1.292	1.664	1.990	2.374	2.639	3.195

Probability less than the critical value ($t_{1-\alpha,\nu}$)						
ν	0.90	0.95	0.975	0.99	0.995	0.999
81.	1.292	1.664	1.990	2.373	2.638	3.194
82.	1.292	1.664	1.989	2.373	2.637	3.193
83.	1.292	1.663	1.989	2.372	2.636	3.191
84.	1.292	1.663	1.989	2.372	2.636	3.190
85.	1.292	1.663	1.988	2.371	2.635	3.189
86.	1.291	1.663	1.988	2.370	2.634	3.188
87.	1.291	1.663	1.988	2.370	2.634	3.187
88.	1.291	1.662	1.987	2.369	2.633	3.185
89.	1.291	1.662	1.987	2.369	2.632	3.184
90.	1.291	1.662	1.987	2.368	2.632	3.183
91.	1.291	1.662	1.986	2.368	2.631	3.182
92.	1.291	1.662	1.986	2.368	2.630	3.181
93.	1.291	1.661	1.986	2.367	2.630	3.180
94.	1.291	1.661	1.986	2.367	2.629	3.179
95.	1.291	1.661	1.985	2.366	2.629	3.178
96.	1.290	1.661	1.985	2.366	2.628	3.177
97.	1.290	1.661	1.985	2.365	2.627	3.176
98.	1.290	1.661	1.984	2.365	2.627	3.175
99.	1.290	1.660	1.984	2.365	2.626	3.175
100.	1.290	1.660	1.984	2.364	2.626	3.174
∞	1.282	1.645	1.960	2.326	2.576	3.090

Confidence Limits of the Mean

- Confidence Limits are a <u>two</u> sided test.
  - i.e. the real mean can be greater than or less then the estimate.
- Example:
  - n = 195   m = 9.2615   s = 0.0228
  - 95% confidence interval
  - α = 0.05
  - $t_{1-\alpha/2,194}$ = 1.9723
  - Lower Limit = m – t * s / sqrt(n) = 9.2615 – 1.9723 * 0.0228 / sqrt(195) = 9.2582
  - Upper Limit = m + t * s / sqrt(n) = 9.2615 + 1.9723 * 0.0228 / sqrt(195) = 9.2647



t distribution in Excel

- Values of t can be obtained using the TINV function in Excel.
- TINV(probability, degrees of freedom)
  - probability = α for a <u>two sided</u> distribution
    - e.g. instead of 1-α/2 = 0.975 in table, use α = 0.05
  - degrees of freedom = ν = n-1
- For the previous example:
  =tinv(0.05, 194) returns 1.9723

alpha	5%
N	195
N - 1 = d of f	194
= TINV(alpha, N-1)	1.9723

- For a one-sided distribution, use 2*α
- Hint: Before using tinv, try duplicating an example in the NIST ESH.

Are two Means Possibly Equal

- We have two estimates of the mean, m_1 and m_2 with $m_1 > m_2$.
- We have two estimates of the standard deviation, s_1 and s_2.
- We have two sample sizes, n_1 and n_2.
- This is a one sided test.
- Null Hypothesis: $\mu_1 = \mu_2$.
- Test Statistic

$$T = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_1^2/n_1 + s_2^2/n_2}}$$

  – Similar to $t = \dfrac{\bar{x} - \mu}{s/\sqrt{n}}$

Are two Means Possibly Equal

- Reject the Null Hypothesis if:

$$T > t_{1-\alpha,\nu}$$

  – where $t_{1-\alpha,\nu}$ is the critical value of the t-distribution with ν degrees of freedom where

$$\nu = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{(s_1^2/n_1)^2/(n_1 - 1) + (s_2^2/n_2)^2/(n_2 - 1)}$$

Equal Variances

- If equal variances are assumed:

$$T = \frac{\bar{x}_1 - \bar{x}_2}{s_p\sqrt{1/n_1 + 1/n_2}}$$

  – where s_p is the pooled estimate of the standard deviation:

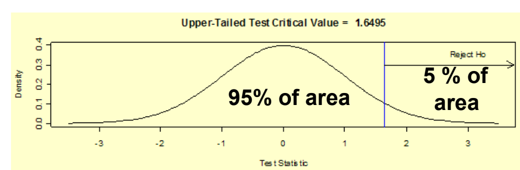
$$s_p = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

- and

$$\nu = n_1 + n_2 - 1$$

Example

- Mileage Data from US and Japanese cars in 1990s
  - $n_1 = 79$ $m_1 = 30.48$ $s_1 = 6.108$
  - $n_2 = 249$ $m_2 = 20.14$ $s_2 = 6.415$
- Assuming variances are equal
  - $T = 12.621$
  - $s_p = 6.343$
  - $\nu = 326$
- For 95% confidence, $\alpha = 0.05$
- $t_{0.95,\nu=326} = 1.6495$
- Since T > t, the hypothesis that the means are equal is rejected!



Large n

- What happens to the confidence limits as n gets large?

$$\mu = \lim_{x \to \infty} (\bar{x} - t_{\alpha,v} \frac{s}{\sqrt{n}}) = \bar{x}$$

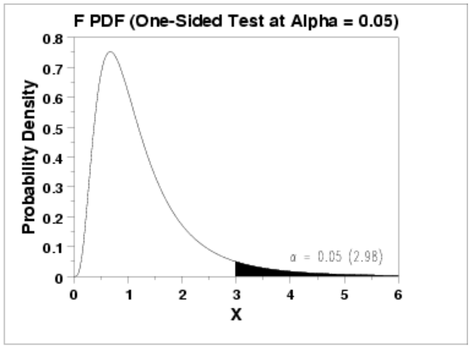
Are Two Variances Equal?

- Null Hypothesis, H0: $\sigma_1^2 = \sigma_2^2$
- Alternative Hypotheses, Ha:
  - $\sigma_1^2 < \sigma_2^2$ for a lower one-tailed test
  - $\sigma_1^2 > \sigma_2^2$ for an upper one-tailed test
  - $\sigma_1^2 \neq \sigma_2^2$ for a two-tailed test
  - Test Statistic: $F = s_1^2/s_2^2$
  - Where s_1^2 and s_2^2 are the sample variances with sample sizes of N_1 and N_2 respectively
  - Significance Level: α

Are Two Variances Equal? (cont.)

- The Hypothesis that the two variances, σ12, and σ22, are equal is rejected if:
  - $F > F_{\alpha,N_1-1,N_2-1}$ for an upper one-tailed test
  - $F < F_{1-\alpha,N_1-1,N_2-1}$ for a lower one-tailed test
  - $F > F_{\alpha,N_1-1,N_2-1}$ or $F < F_{1-\alpha,N_1-1,N_2-1}$ for a two-tailed test
- where F_{α,N_1-1,N_2-1} is the critical value of the F distribution with N_1-1 and N_2-1 degrees of freedom and a significance level of α.

F Distribution



F Distribution

Upper critical values of the F distribution

for ν_1 numerator degrees of freedom and ν_2 denominator degrees of freedom

5% significance level

$F_{.05}(\nu_1, \nu_2)$

ν_2 \ ν_1	1	2	3	4	5	6	7	8	9	10
1	161.448	199.500	215.707	224.583	230.162	233.986	236.768	238.882	240.543	241.882
2	18.513	19.000	19.164	19.247	19.296	19.330	19.353	19.371	19.385	19.396
3	10.128	9.552	9.277	9.117	9.013	8.941	8.887	8.845	8.812	8.786
4	7.709	6.944	6.591	6.388	6.256	6.163	6.094	6.041	5.999	5.964
5	6.608	5.786	5.409	5.192	5.050	4.950	4.876	4.818	4.772	4.735
6	5.987	5.143	4.757	4.534	4.387	4.284	4.207	4.147	4.099	4.060
7	5.591	4.737	4.347	4.120	3.972	3.866	3.787	3.726	3.677	3.637
8	5.318	4.459	4.066	3.838	3.687	3.581	3.500	3.438	3.388	3.347
9	5.117	4.256	3.863	3.633	3.482	3.374	3.293	3.230	3.179	3.137
10	4.965	4.103	3.708	3.478	3.326	3.217	3.135	3.072	3.020	2.978
11	4.844	3.982	3.587	3.357	3.204	3.095	3.012	2.948	2.896	2.854
12	4.747	3.885	3.490	3.259	3.106	2.996	2.913	2.849	2.796	2.753
13	4.667	3.806	3.411	3.179	3.025	2.915	2.832	2.767	2.714	2.671
14	4.600	3.739	3.344	3.112	2.958	2.848	2.764	2.699	2.646	2.602
15	4.543	3.682	3.287	3.056	2.901	2.790	2.707	2.641	2.588	2.544
16	4.494	3.634	3.239	3.007	2.852	2.741	2.657	2.591	2.538	2.494
17	4.451	3.592	3.197	2.965	2.810	2.699	2.614	2.548	2.494	2.450
18	4.414	3.555	3.160	2.928	2.773	2.661	2.577	2.510	2.456	2.412
19	4.381	3.522	3.127	2.895	2.740	2.628	2.544	2.477	2.423	2.378
20	4.351	3.493	3.098	2.866	2.711	2.599	2.514	2.447	2.393	2.348
21	4.325	3.467	3.072	2.840	2.685	2.573	2.488	2.420	2.366	2.321
22	4.301	3.443	3.049	2.817	2.661	2.549	2.464	2.397	2.342	2.297
23	4.279	3.422	3.028	2.796	2.640	2.528	2.442	2.375	2.320	2.275
24	4.260	3.403	3.009	2.776	2.621	2.508	2.423	2.355	2.300	2.255
25	4.242	3.385	2.991	2.759	2.603	2.490	2.405	2.337	2.282	2.236

Using Excel for F Dist

- Use finv(α, N-1, N-1) function in Excel.
  - Use exampel in NIST ESH to check usage.

Alpha	0.05
N1	240
N2	240
FINV(a/2, N1-1, N2-2)=	1.289384
FINV(1-a/2, N1-1, N2-1)=	0.775564

Ceramic Data Example

```
BATCH 1:
   NUMBER OF OBSERVATIONS         =      240
   MEAN                           =    688.9987
   STANDARD DEVIATION             =    65.54909

BATCH 2:
   NUMBER OF OBSERVATIONS         =      240
   MEAN                           =    611.1559
   STANDARD DEVIATION             =    61.85425
```

Is 65.5 significantly different from 61.9?

H_0: $\sigma_1^2 = \sigma_2^2$

H_a: $\sigma_1^2 \neq \sigma_2^2$

Test statistic: $F = 1.123037$

Numerator degrees of freedom: $N_1 - 1 = 239$

Denominator degrees of freedom: $N_2 - 1 = 239$

Significance level: $\alpha = 0.05$

Critical values: $F(1-\alpha/2, N_1-1, N_2-1) = 0.7756$

$F(\alpha/2, N_1-1, N_2-1) = 1.2894$

Rejection region: Reject H_0 if $F < 0.7756$ or $F > 1.2894$

From Excel

Alpha	0.05
N1	240
N2	240
FINV(a/2, N1-1, N2-2)=	1.289384
FINV(1-a/2, N1-1, N2-1)=	0.775564