

平成25年度AIRIS

## 統計の基礎(3)

・2つのデータを比較する

平成25年9月27日

防衛医学研究センター 感染症疫学対策研究官

教授 加來浩器 (KAKU KOKI)

### 2つの群のデータ比較の流れ

**A群**

平均  
不偏分散  
標準偏差  
信頼区間

**B群**

平均  
不偏分散  
標準偏差  
信頼区間

それぞれの群の代表値を求める

両群の平均値の差、差の分散、差の標本誤差を求める

t分布表を用いて両群の平均値の差の信頼区間を求める。

A中学校1年男子の2つのクラスの身長を代表値を求める

組-番号	身長
1-1	137.0
1-2	144.0
1-3	149.0
1-4	136.0
1-5	142.0
1-6	156.0
1-7	150.0
1-8	138.0
1-9	146.0
1-10	157.0
1-11	160.0
1-12	153.0
1-13	149.0
1-14	142.0
1-15	147.0
1-16	135.0

組-番号	身長
2-1	144
2-2	143
2-3	139
2-4	166
2-5	169
2-6	144
2-7	147
2-8	138
2-9	176
2-10	133
2-11	170
2-12	137
2-13	146
2-14	140
2-15	122
2-16	177

データの数、平均、不偏分散、標準偏差を求める

各組の標準誤差を計算する

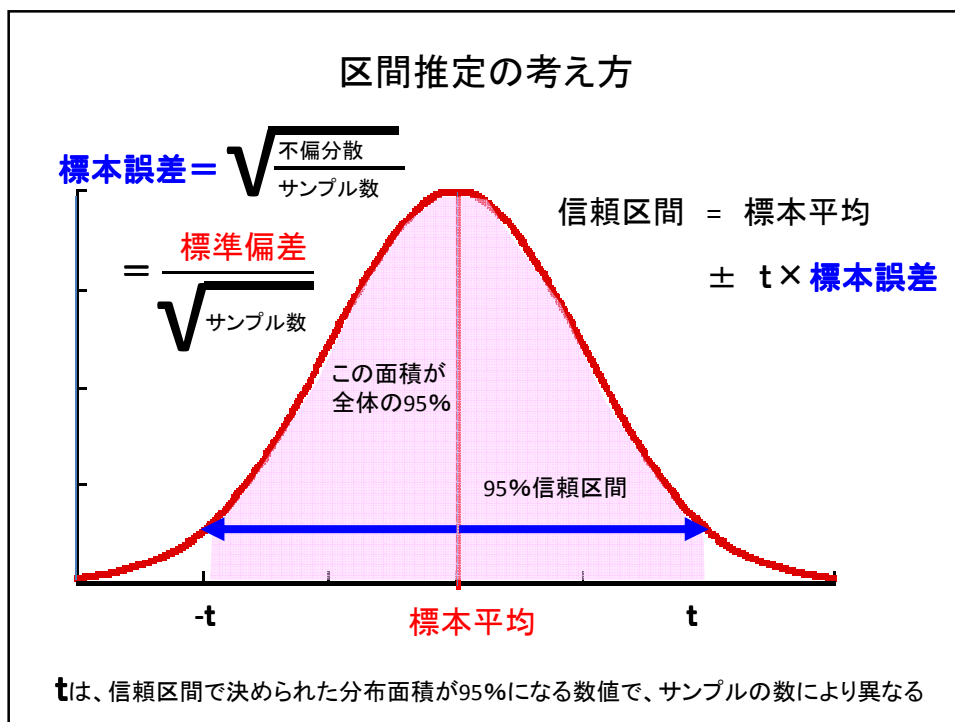
The screenshot shows an Excel spreadsheet with two columns of student height data. The first column is for Class 1 (1-1 to 1-16) and the second is for Class 2 (2-1 to 2-16). Below the data, there are formulas for calculating the number of data points (COUNT), average (AVERAGE), unbiased variance (VAR), and standard deviation (STDEV). The formulas are: COUNT, AVERAGE, VAR, and STDEV. The spreadsheet also shows the results of these calculations for each class.

**自由度は、(データの数-1)で示される**

**自由度のt値をt分布表から読み取る**

**95%信頼区間の下限、上限を求める**

自由度	95%	90%	80%	70%	60%	50%	40%	30%	20%	10%
1	63.657	63.657	63.657	63.657	63.657	63.657	63.657	63.657	63.657	63.657
2	4.303	3.982	3.747	3.581	3.465	3.385	3.326	3.281	3.245	3.216
3	3.182	2.924	2.747	2.591	2.500	2.438	2.398	2.363	2.333	2.309
4	2.776	2.576	2.423	2.291	2.214	2.157	2.121	2.090	2.064	2.041
5	2.576	2.377	2.224	2.093	2.017	1.960	1.924	1.893	1.867	1.845
6	2.447	2.248	2.095	1.964	1.888	1.831	1.795	1.764	1.738	1.716
7	2.365	2.166	2.013	1.882	1.806	1.749	1.713	1.682	1.656	1.634
8	2.306	2.107	1.954	1.823	1.747	1.690	1.654	1.623	1.597	1.575
9	2.262	2.063	1.910	1.779	1.703	1.646	1.610	1.579	1.553	1.531
10	2.228	2.029	1.876	1.745	1.669	1.612	1.576	1.545	1.519	1.497
11	2.201	2.002	1.849	1.718	1.642	1.585	1.549	1.518	1.492	1.470
12	2.177	1.978	1.825	1.694	1.618	1.561	1.525	1.494	1.468	1.446
13	2.157	1.958	1.805	1.674	1.598	1.541	1.505	1.474	1.448	1.426
14	2.140	1.943	1.790	1.659	1.583	1.526	1.490	1.459	1.433	1.411
15	2.131	1.934	1.781	1.650	1.574	1.517	1.481	1.450	1.424	1.402
16	2.124	1.928	1.775	1.644	1.568	1.511	1.475	1.444	1.418	1.396
17	2.118	1.923	1.770	1.639	1.563	1.506	1.470	1.439	1.413	1.391
18	2.113	1.919	1.766	1.635	1.559	1.502	1.466	1.435	1.409	1.387
19	2.109	1.916	1.763	1.632	1.556	1.499	1.463	1.432	1.406	1.384
20	2.106	1.914	1.761	1.630	1.554	1.497	1.461	1.430	1.404	1.382
21	2.104	1.913	1.760	1.629	1.553	1.496	1.460	1.429	1.403	1.381
22	2.102	1.912	1.759	1.628	1.552	1.495	1.459	1.428	1.402	1.380
23	2.101	1.911	1.758	1.627	1.551	1.494	1.458	1.427	1.401	1.379
24	2.100	1.910	1.757	1.626	1.550	1.493	1.457	1.426	1.400	1.378
25	2.099	1.910	1.756	1.625	1.549	1.492	1.456	1.425	1.399	1.377
26	2.098	1.909	1.755	1.624	1.548	1.491	1.455	1.424	1.398	1.376
27	2.098	1.909	1.755	1.624	1.548	1.491	1.455	1.424	1.398	1.376
28	2.097	1.908	1.754	1.623	1.547	1.490	1.454	1.423	1.397	1.375
29	2.097	1.908	1.754	1.623	1.547	1.490	1.454	1.423	1.397	1.375
30	2.097	1.908	1.754	1.623	1.547	1.490	1.454	1.423	1.397	1.375



関数: 各グループの両群の身長データの信頼区間を算出する。

A群の身長データの身長		B群の身長データの身長		人口有意差	
身長	有意差	身長	有意差	身長	有意差
1-1	127.0	1-1	144	1	3100
1-2	144.0	1-2	140	2	3110
1-3	148.0	1-3	133	3	3115
1-4	136.0	1-4	169	4	3121
1-5	140.0	1-5	160	5	3147
1-6	152.0	1-6	144	6	3200
1-7	152.0	1-7	147	7	3206
1-8	150.0	1-8	155	8	3210
1-9	148.0	1-9	176	9	3224
1-10	157.0	1-10	133	10	3207
1-11	148.0	1-11	170	11	3179
1-12	153.0	1-12	137	12	3169
1-13	148.0	1-13	146	13	3140
1-14	140.0	1-14	140	14	3137
1-15	147.0	1-15	132	15	3133
1-16	130.0	1-16	177	16	3110
17		17		17	3107
18		18		18	3080
19		19		19	3080
20		20		20	3089
21		21		21	3089
22		22		22	3089
23		23		23	3089
24		24		24	3084
25		25		25	3089
26		26		26	3084
27		27		27	3089
28		28		28	3084
29		29		29	3040
30		30		30	3040
31		31		31	3040
32		32		32	3040

信頼区間を「両群の平均の差」に適用してみる

「両群の平均の差」の信頼区間

$$= (\text{標本平均A} - \text{標本平均B}) \pm t \times \text{差の標準誤差}$$

**平均の差、差の標準誤差を計算する**

**差の標準誤差 =**  
 (1組の不偏分散 / データ数 + 2組の不偏分散 / データ数) の平方根

組別	値	組別	値
1-1	127.8	2-1	144
1-2	144.8	2-2	187
1-3	148.8	2-3	239
1-4	129.8	2-4	189
1-5	142.8	2-5	189
1-6	126.8	2-6	144
1-7	122.8	2-7	147
1-8	128.8	2-8	178
1-9	148.8	2-9	179
1-10	127.8	2-10	123
1-11	168.8	2-11	170
1-12	120.8	2-12	127
1-13	148.8	2-13	146
1-14	142.8	2-14	140
1-15	147.8	2-15	122
1-16	125.8	2-16	117

**自由度 =**  
 (1組データ数 - 1) + (2組データ数 - 1) で示される

**自由度のt値をt分布表から読み取る**

統計量	値
データ数	18
平均	127.2
不偏分散	80.1
平均の差	-2.1
差の標準誤差	4.80

**95%信頼区間の下限、上限を求める**

身長	数値	身長	数値
171	127.2	171	144
172	144.0	172	187
173	148.0	173	139
174	129.0	174	180
175	142.0	175	180
176	126.0	176	144
177	129.0	177	147
178	128.0	178	178
179	148.0	179	170
180	127.0	180	123
181	168.0	181	150
182	120.0	182	127
183	148.0	183	140
184	142.0	184	140
185	147.0	185	122
186	120.0	186	117

項目	値	項目	値
グループの数	18	グループの数	18
平均	142.2	平均	142.2
分散	80.1	分散	215.0
標準偏差	8.95	標準偏差	14.66
信頼区間下限	128.1	信頼区間下限	128.1
信頼区間上限	156.3	信頼区間上限	156.3

**1組と2組の男子生徒の身長差の信頼区間が、( マイナス )から( プラス )に分布していることは、平均値に有意な差が( 無い )ことを意味する。**