

	正規分布		その他の分布
	σ既知	σ未知	
$n \geq 100$	Z検定	Z検定	Z検定
$n < 100$	Z検定	t検定	ノンパラメトリック検定

正規分布
N数が十分多い →

↑

Shapiro-Wilk (シャピロ-ウィルク) 検定

↓

正規分布でない
母集団分布不明 →
N数が少ない

パラメトリック検定	Z検定	平均 比率
	t検定	
ノンパラメトリック検定	χ ² 検定	度数 ・適合度の検定 ・独立性の検定
		母分散の検定
	U検定 (対応有り) サインランク検定 (対応なし)	平均順位

$n \geq 100$ Z検定 母集団の標準偏差未知

東京都に住む20歳以上の人の中から

男性250人、女性200人をランダムに選び、

電化製品を5点満点で評価してもらいました。

その結果、男性は平均点2.72点、標準偏差1.00点

女性は平均点3.39点、標準偏差0.95点

評価に差があるといえるかどうか、有意水準0.05で検定しなさい

帰無仮説：男性と女性の評価は同じ

標本サイズ $n_1 = 250$ 、 $n_2 = 200$

標本平均 $\bar{x}_1 = 2.72$ 、 $\bar{x}_2 = 3.39$

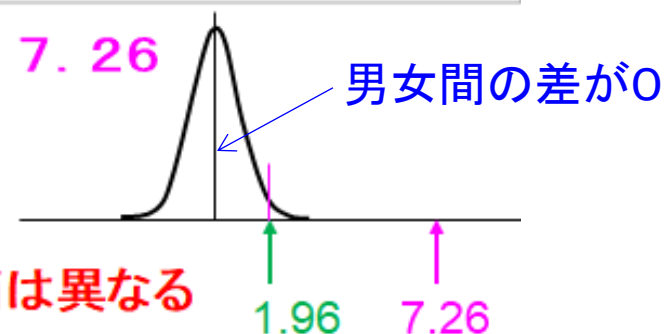
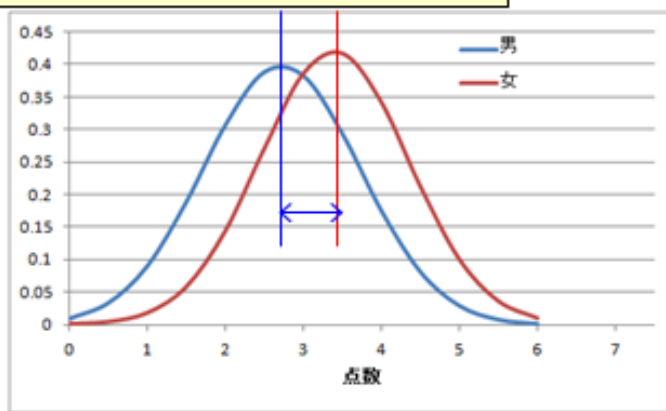
標本標準偏差 $u_1 = 1.00$ 、 $u_2 = 0.95$

統計量

$$T = \frac{\bar{x}_2 - \bar{x}_1}{\sqrt{\frac{u_1^2}{n_1} + \frac{u_2^2}{n_2}}} = \frac{3.39 - 2.72}{\sqrt{\frac{1}{250} + \frac{0.9025}{200}}} = 7.26$$

棄却域 両側検定 $Z(0.025) = 1.96$

比較 $|T| = 7.26 > 1.96$ 結論 男女の評価は異なる



棄却域限界のZ値

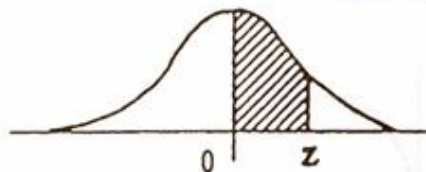
付表(1) 標準正規分布表

有意水準0.05(5%)=信頼度95%

片側検定 $50 - 5 = 45\%$ → 1.64

両側検定 $50 - 5 / 2 = 47.5\%$ → 1.96

値Zに対する確率



Excel関数

NORMSINV(0.975)=1.96

1σの確率

0.34134×2

=0.68268

2σの確率

0.47725×2

=0.9545

3σの確率

0.49865×2

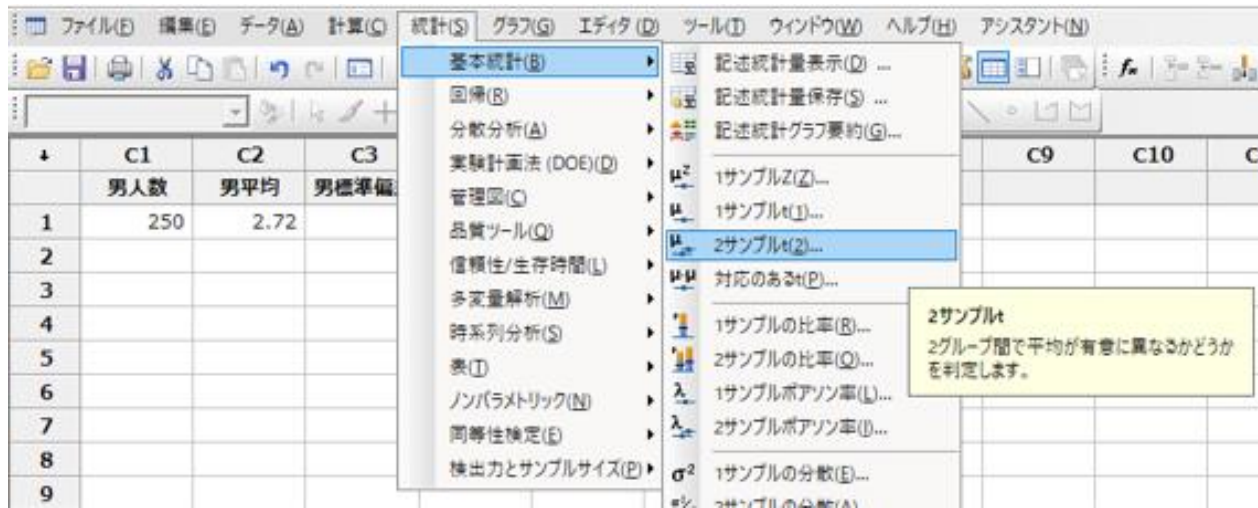
=0.9973

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.00000	0.00399	0.00798	0.01197	0.01595	0.01994	0.02392	0.02790	0.03188	0.03586
0.1	.03983	.04380	.04776	.05172	.05567	.05962	.06356	.06749	.07142	.07535
0.2	.07926	.08317	.08706	.09095	.09483	.09871	.10257	.10642	.11026	.11409
0.3	.11791	.12172	.12552	.12930	.13307	.13683	.14058	.14431	.14803	.15173
0.4	.15542	.15910	.16276	.16640	.17003	.17364	.17724	.18082	.18439	.18793
0.5	.19146	.19497	.19847	.20194	.20540	.20884	.21226	.21566	.21904	.22240
0.6	.22575	.22907	.23237	.23565	.23891	.24215	.24537	.24857	.25175	.25490
0.7	.25804	.26115	.26424	.26730	.27035	.27337	.27637	.27935	.28230	.28524
0.8	.28814	.29103	.29389	.29673	.29955	.30234	.30511	.30785	.31057	.31327
0.9	.31594	.31859	.32121	.32381	.32639	.32894	.33147	.33398	.33646	.33891
1.0	.34134	.34375	.34614	.34850	.35083	.35314	.35543	.35769	.35993	.36214
1.1	.36433	.36650	.36864	.37076	.37286	.37493	.37698	.37900	.38100	.38298
1.2	.38493	.38686	.38877	.39065	.39251	.39435	.39617	.39796	.39973	.40147
1.3	.40320	.40490	.40658	.40824	.40988	.41149	.41309	.41466	.41621	.41774
1.4	.41924	.42073	.42220	.42364	.42507	.42647	.42786	.42922	.43056	.43189
1.5	.43319	.43448	.43574	.43699	.43822	.43948	.44062	.44179	.44295	.44408
1.6	.44520	.44630	.44738	.44845	.44950	.45058	.45154	.45254	.45352	.45449
1.7	.45543	.45637	.45728	.45818	.45907	.45994	.46080	.46164	.46246	.46327
1.8	.46407	.46485	.46562	.46638	.46712	.46784	.46856	.46926	.46995	.47062
1.9	.47128	.47193	.47257	.47320	.47381	.47441	.47500	.47558	.47615	.47670
2.0	.47725	.47778	.47831	.47882	.47932	.47982	.48030	.48077	.48124	.48169
2.1	.48214	.48257	.48300	.48341	.48382	.48422	.48461	.48500	.48537	.48574
2.2	.48610	.48645	.48679	.48713	.48745	.48778	.48809	.48840	.48870	.48899
2.3	.48928	.48956	.48983	.49010	.49036	.49061	.49086	.49111	.49134	.49158
2.4	.49180	.49202	.49224	.49245	.49266	.49286	.49305	.49324	.49343	.49361
2.5	.49379	.49396	.49413	.49430	.49446	.49461	.49477	.49492	.49506	.49520
2.6	.49534	.49547	.49560	.49573	.49585	.49598	.49609	.49621	.49632	.49643
2.7	.49653	.49664	.49674	.49683	.49693	.49702	.49711	.49720	.49728	.49736
2.8	.49744	.49752	.49760	.49767	.49774	.49781	.49788	.49795	.49801	.49807
2.9	.49813	.49819	.49825	.49831	.49836	.49841	.49846	.49851	.49856	.49861
3.0	.49865	.49869	.49874	.49878	.49882	.49886	.49889	.49893	.49897	.49900
3.1	.49903	.49906	.49910	.49913	.49916	.49918	.49921	.49924	.49926	.49929

MinitabによるZ検定

2サンプルt(2)と書いてあり、Z検定は見当たらない

①統計→基本統計→2サンプルt(2)

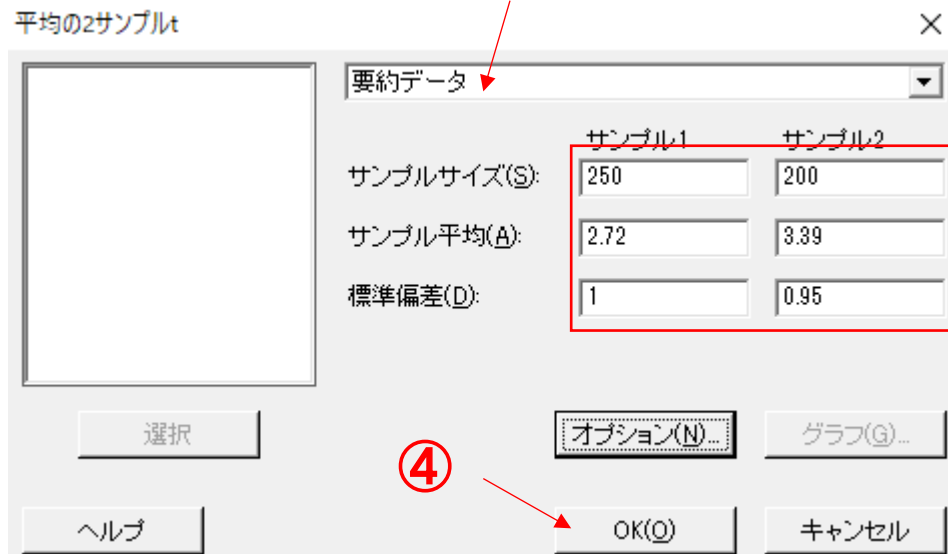


実行結果

2サンプルt検定と信頼区間

サン プル	非欠 損値	平均	標準偏差	平均の標 準誤差
1	250	2.72	1.00	0.063
2	200	3.390	0.950	0.067

②要約データ 選択



③データ入力

差 = $\mu(1) - \mu(2)$
差を推定: -0.6700
差に対する95%の信頼区間: (-0.8513, -0.4887)
差=0 (対 ≠) のT検定: t値=-7.26 p値=0.000 DF=434

統計量t=7.26であるが
閾値の1.96がないので、
判定不能

P値=0.000<0.05なので
有意差あり