

統計演習 (14)

χ^2 検定

高谷 智英

Jan 17, 2018





χ^2 検定 (Chi-squared test)

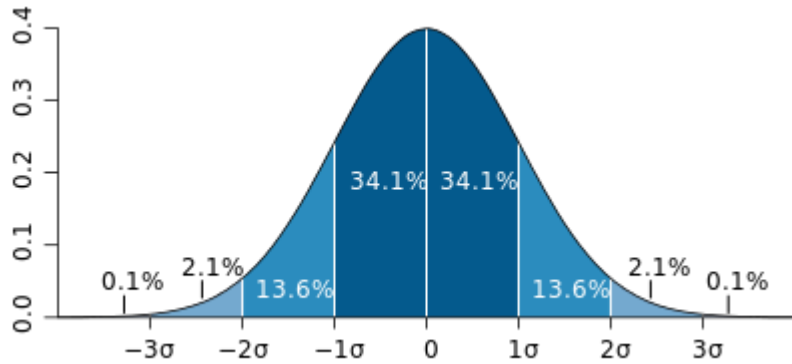
χ^2 検定 : データが **χ^2 分布** に従っているかどうかを検定する。

χ^2 分布 : 標準正規分布 (p34) に従う、
k 個の確率変数 (X_1, X_2, \dots, X_k) の
二乗の和 ($X_1^2 + X_2^2 + \dots + X_k^2$) が従う分布を
自由度 k の χ^2 分布という。

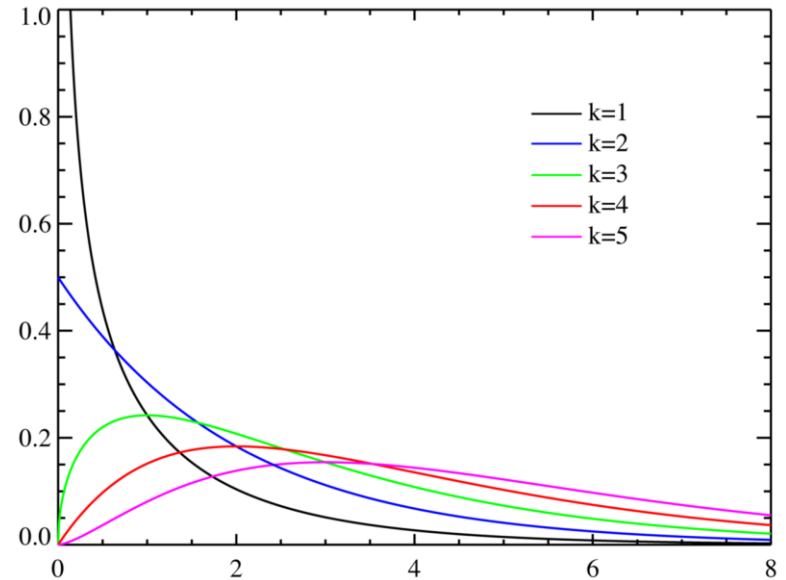
データの分布が期待される頻度に比べて偏っているかを検定する。



標準正規分布と χ^2 分布



標準正規分布



自由度 k (1~5) の χ^2 分布

データが標準正規分布に従っているなら、 χ^2 の和が大きくなる確率は少ないはずである。



χ^2 検定：期待値が理論的に明らかな場合

帰無仮説：データの分布は期待値に適合している。

帰無仮説が棄却される ($p < 0.05$, etc) と...
データの分布に偏りがある

Excel の関数で、 χ^2 検定の p 値を求めることができる。

p 値 = CHITEST (実測値, 期待値)



χ^2 検定：例題 1（メンデル遺伝）

花色が赤く草丈が高い植物個体と、
花色が白く草丈が低い植物個体を交配して F1 を得た。

F1 は全て花色が赤く草丈が高い個体であった。

F1 どうしを交配した結果、F2 の表現型は以下となった。

F2表現型	個体数
赤・高	65
赤・低	50
白・高	30
白・低	15
計	160



χ^2 検定：例題 1（メンデル遺伝）

赤花・白花を決定する遺伝子の遺伝子座が同じで、
赤花の対立遺伝子が優性であり、

高花・低花を決定する遺伝子の遺伝子座が同じで、
高花の対立遺伝子が優性であるなら、

F2 表現型の分離比は **9:3:3:1** になる。（作業仮説）

F1配偶子の 遺伝子型	赤・高	赤・低	白・高	白・低
赤・高	赤赤・高高 赤・高	赤赤・高低 赤・高	赤白・高高 赤・高	赤白・高低 赤・高
赤・低	赤赤・高低 赤・高	赤赤・低低 赤・低	赤白・高低 赤・高	赤白・低低 赤・低
白・高	赤白・高高 赤・高	赤白・高低 赤・高	白白・高高 白・高	白白・高低 白・高
白・低	赤白・高低 赤・高	赤白・低低 赤・低	白白・高低 白・高	白白・低低 白・低



χ^2 検定：例題 1（メンデル遺伝）

作業仮説が正しいければ、F2 の表現型の期待値は次のようになる。

F2表現型	個体数	分離比	期待値
赤・高	65	9	90
赤・低	50	3	30
白・高	30	3	30
白・低	15	1	10
計	160	16	160

帰無仮説：実測値（個体数）の分布は期待値に適合している。
＝ 表現型の分離比は 9:3:3:1 である。



χ^2 検定：例題 1（メンデル遺伝）

B8		fx		=CHITEST(B2:B5,D2:D5)		
	A	B	C	D	E	F
1	表現型	個体数	分離比	期待値		
2	赤・高	65	9	90		
3	赤・低	50	3	30		
4	白・高	30	3	30		
5	白・低	15	1	10		
6	計	160	16	160		
7						
8	p値	4.49E-05				
9						
10						



χ^2 検定：例題 1（メンデル遺伝）

F2表現型	個体数	分離比	期待値
赤・高	65	9	90
赤・低	50	3	30
白・高	30	3	30
白・低	15	1	10
計	160	16	160

帰無仮説：実測値（個体数）の分布は期待値に適合している。
＝ 表現型の分離比は 9:3:3:1 である。

χ^2 検定の p 値 = 4.49×10^{-5}

であるから、帰無仮説は棄却される。

表現型の分離比は 9:3:3:1 ではない。作業仮説は否定される。



χ^2 検定：例題 2（メンデル遺伝）

トウモロコシのF₂の表現型について、以下の結果を得た。
表現型の分離比が 9:3:3:1 であるかを検定せよ。

表現型	個体数	分離比	期待値
緑・無地	773	9	
金・無地	231	3	
緑・縞	238	3	
金・縞	59	1	
計		16	



χ^2 検定：期待値が必ずしも理論的に明らかでない場合

例：アンケートの回答結果など

S大生N学部D物コースの学生に尋ねました。
どの先生の研究室に配属されたいですか？

研究室	男子	女子
S里先生	19	32
Y倉先生	31	17
T中先生	11	44
T谷先生	2	5
計	63	98



χ^2 検定：アンケートの分析（1）

期待値の分布は必ずしも明らかではないが、回答結果が均等な分布からズれているかどうかを検定することは可能である。

研究室	男子	分離比	期待値
S里先生	19	25%	15.75
Y倉先生	31	25%	15.75
T中先生	11	25%	15.75
T谷先生	2	25%	15.75
計	63	100%	63

帰無仮説：男子の配属希望は各研究室で均等である、
を検定せよ。



χ^2 検定：アンケートの分析 (2)

男女で回答パターンに違いがあるか？

実測値

研究室	男子	女子	計
S里先生	19	32	51
Y倉先生	31	17	48
T中先生	11	44	55
T谷先生	2	5	7
計	63	98	161

期待値

研究室	男子	女子	計
S里先生	20.0	31.0	51
Y倉先生	18.8	29.2	48
T中先生	21.5	33.5	55
T谷先生	2.7	4.3	7
計	63	98	161