AS統計科学　適合度と独立性の検定

【適合度の検定】

母集団が k 個の排反な事象 A_1，A_2，…，A_k に分けられており，A_1，A_2，…，A_k の現れる確率（理論値）をそれぞれ p_1，p_2，…，p_k とする。母集団から大きさ n の標本を無作為抽出したとき，事象 A_1，A_2，…，A_k の観測度数がそれぞれ f_1，f_2，…，$f_k\left(\sum_{i=1}^{k} f_i = n\right)$ であったとする。このとき，観測度数が理論値から得られる期待度数と適合しているかを検定したい。

そこで，次のように帰無仮説 H_0，対立仮説 H_1 を定め，有意水準 α で検定する。

H_0：観測度数と期待度数は等しい（理論と実験の結果は適合している）

H_1：観測度数と期待度数は等しくない（理論と実験の結果は適合していない）

帰無仮説 H_0 のもとでは，期待度数は標本の大きさと確率の積 np_i で表される。

事象	A_1	A_2	$\cdots$	A_k	計
観測度数	f_1	f_2	$\cdots$	f_k	n
期待度数	np_1	np_2	$\cdots$	np_k	n

このとき，事象 A_i が現れる回数を確率変数 X_i で表すと，$\displaystyle\sum_{i=1}^{k}\frac{(X_i-np_i)^2}{np_i}$ は，$i=1,2,\cdots,k$ に対し $np_i \geqq 5$ であるとき自由度 $df=k-1$ の χ^2 分布に従う。よって，χ^2 の実現値 $\displaystyle\chi^2{}_{*}=\sum_{i=1}^{k}\frac{(f_i-np_i)^2}{np_i}$ を計算し，

$\chi^2{}_{*} \geqq \chi^2{}_{(k-1)}(\alpha)$ ならば H_0 を棄却し，$\chi^2{}_{*} < \chi^2{}_{(k-1)}(\alpha)$ ならば H_0 を採択する。

例題１

メンデルの遺伝法則によれば，えんどう豆の交配実験の結果は，円形黄色，角形黄色，円形緑色，角形緑色の現れる割合が 9:3:3:1 になるといわれている。この交配実験を行ってみたところ，次の数値を得た。

種子の種類	円形黄色	角形黄色	円形緑色	角形緑色	計
観測度数	195	49	64	12	320

このとき，この実験結果は理論に適合しているといえるだろうか。有意水準5%で検定せよ。

【独立性の検定】

母集団が２つの属性A，Ｂにより，それぞれの排反事象 A_1， A，…，A_k および B， B_2，…， B_l に分類されているものとする。この母集団から n 個の標本を無作為抽出したとき，次のような観測度数の表ができたとする。これを $k\times l$ 分割表という。

	B_1	B_2	$\cdots$	B_l	計
A_1	f_{11}	f_{12}	$\cdots$	f_{1l}	f_{1*}
A_2	f_{21}	f_{22}	$\cdots$	f_{2l}	f_{2*}
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
A_k	f_{k1}	f_{k2}	$\cdots$	f_{kl}	f_{k*}
計	f_{*1}	f_{*2}	$\cdots$	f_{*l}	n

このとき，属性Ａと属性Ｂが独立であることを検定したい。

そこで，次のように帰無仮説 H_0，対立仮説 H_1 を定め，有意水準 α で検定する。

H_0 ：属性Ａと属性Ｂは独立である（ＡとＢは無関係である）

H_1 ：属性Ａと属性Ｂは独立でない（ＡとＢは無関係ではない）

f_{ij} に対応する確率変数を X_{ij} とすると， H_0 が真であれば， $\chi^2=\sum_{i=1}^{k}\sum_{j=1}^{l}\dfrac{\left(X_{ij}-\dfrac{f_{i*}\times f_{*j}}{n}\right)^2}{\dfrac{f_{i*}\times f_{*j}}{n}}$ は，すべての i， j に対して $\dfrac{f_{i*}\times f_{*j}}{n}\geqq 5$ のとき，近似的に自由度 $df=(k-1)(l-1)$ の χ^2 分布に従う。よって， χ^2 の実現値 ${\chi^2}_*$ が，

${\chi^2}_*=\sum_{i=1}^{k}\sum_{j=1}^{l}\dfrac{\left(f_{ij}-\dfrac{f_{i*}\times f_{*j}}{n}\right)^2}{\dfrac{f_{i*}\times f_{*j}}{n}}\geqq\chi^2_{(k-1)(l-1)}(\alpha)$ であれば H_0 を棄却し， ${\chi^2}_*<x^2_{(k-1)(l-1)}(\alpha)$ ならば H_0 を採択する。

例題２

ある新薬の効果を検証するために，100人の患者について臨床検査を行ったところ，次の結果を得た。

	治った患者	治らなかった患者	計
新薬を投与された患者	45	15	60
偽薬を投与された患者	20	20	40
計	65	35	100

この新薬は効果があったのか，有意水準5%で検定せよ。