§8.1　仮説検定

例として，表側と裏側とがある円盤を回して倒して表側が上を向くか裏側が上を向くかを調べる試行において，裏側が上を向く確率が $\frac{1}{2}$ なのかどうかを考えます．例えば，この試行を独立に 8 回繰り返して，8 回とも裏側が上を向いたとします．仮に裏側が上を向く確率が $\frac{1}{2}$ であるとしてみると，8 回とも裏が上を向くということは，あり得ないことではありませんがその確率はごく小さいものです；つまり確率的にひどく不自然です．ですから裏側が上を向く確率が $\frac{1}{2}$ であるとは考えにくいといえます．そこでこのようなときは裏側が上を向く確率は $\frac{1}{2}$ でない，裏側が上を向く確率は表側が上を向く確率より大きい，と一応判断する — というのがこれから述べる仮説検定の基本的な考え方です．つまり，仮説検定では次のように考えます：ある仮定の下で確率的にひどく不自然なことが起きたとき，その仮定は成り立たないと一応判断する．

もう少し具体的な例を述べます．実数 m と s ($s>0$) とが与えられているとします．統計量 X は正規分布に従うとします．X の平均値と分散とは分かりません．この統計量 X について次の仮説 H を立てます：

$$\mathrm{H}\ :\ \mathrm{E}[X]=m \ \text{かつ}\ \mathrm{V}[X]=s^2 \ .$$

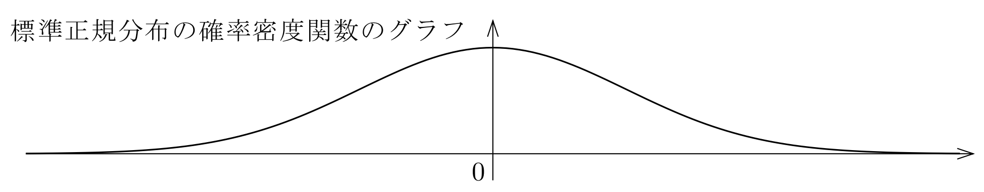
この仮説 H が本当かどうかは分かりません．仮説 H が妥当かどうかどうかを判断する基準を作ります．そして，仮説 H を仮定したとき確率的に不自然なことが起きるならば，仮説 H は成り立たないと判断する．

仮説 H を仮定すると，確率変数 X は正規分布 $\mathrm{N}(m,s^2)$ に従うので，X を標準化した確率変数 $T=\dfrac{X-m}{s}$ は標準正規分布 $\mathrm{N}(0,1)$ に従います．標準正規分布 $\mathrm{N}(0,1)$ に従う確率変数 T の累積分布関数を $\Phi(x)$ とおきます：

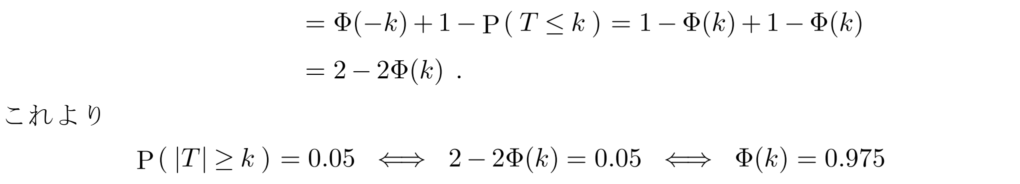
$$\Phi(x) = \mathrm{P}(\,T\le x\,) = \int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{t^2}{2}\right)dt \ .$$

各実数 a について，

$$\mathrm{P}(\,Z\le a\,)=\Phi(a)\ ,$$
$$\mathrm{P}(\,Z\ge a\,)=1-\mathrm{P}(\,Z\le a\,)=1-\Phi(a)\ ,$$
$$\Phi(-a)=1-\Phi(a)\ .$$



確率変数 T が標準正規分布に従うとき，T の実現値は，0 に近い実数になる確率が大きく，0 からかけ離れた実数になる確率は小さいです．ですから，T の実現値が 0 から離れるほど，確率的に仮説 H は疑わしくなります．そこで，T の実現値が 0 からかけ離れた値になるとき，仮説 H は成り立たないと判断します．ここで，T の実現値について " 0 からかけ離れている " という範囲をどう定めるかが問題になります．そこで次のような正の定数 k を考えます：T の実現値が 0 から k 以上離れれる確率が 5% になる，つまり $\mathrm{P}(\,|T|\ge k\,)=0.05$ ．



仮説 H より確率変数 T は標準正規分布 $\mathrm{N}(0,1)$ に従うので $\mathrm{P}(\,T\le x\,)=\Phi(x)$ ．

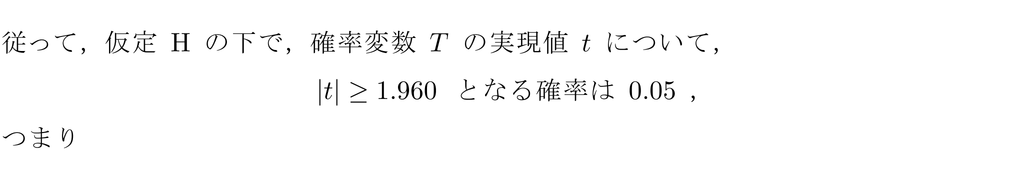
$$\begin{aligned}\mathrm{P}(\,|T|\ge k\,) &= \mathrm{P}(\,T\le -k\,)+\mathrm{P}(\,T\ge k\,)\\ &=\Phi(-k)+1-\mathrm{P}(\,T\le k\,)=1-\Phi(k)+1-\Phi(k)\\ &=2-2\Phi(k)\ .\end{aligned}$$

これより

$$\begin{aligned}\mathrm{P}(\,|T|\ge k\,)=0.05 &\iff 2-2\Phi(k)=0.05 \iff \Phi(k)=0.975\\ &\iff k=\Phi^{-1}(0.975)\ .\end{aligned}$$

$\Phi^{-1}(0.975)\fallingdotseq 1.960$ ですから，

$$\mathrm{P}(\,|T|\ge 1.960\,) = 0.05\ .$$



仮説 H を仮定すると，統計量 T は標準正規分布 $\mathrm{N}(0,1)$ に従いますから，

$$\mathrm{P}(\,|T|\ge 1.960\,) = 0.05\ .$$

従って，仮定 H の下で，確率変数 T の実現値 t について，

$$|t|\ge 1.960 \ \text{となる確率は}\ 0.05\ ,$$

つまり

$$t \ \text{の値が}\ 0\ \text{から}\ 1.960\ \text{以上離れる確率は}\ 0.05$$

です．この確率 $0.05=5\%$ は極めて小さいと考えて，実際に $|t|\ge 1.960$ となるとき，H を仮定したことに無理があったと推測します．つまり，$|t|\ge 1.960$ となるとき，仮説 H は成り立たないと判断します．

実数 m と s ($s>0$) とが与えられているとします．正規分布に従う統計量 X について，仮説 " $\mathrm{E}[X]=m$ かつ $\mathrm{V}[X]=s^2$ " が疑わしいといえるかどうか考えます．そこで，仮説 " $\mathrm{E}[X]=m$ かつ $\mathrm{V}[X]=s^2$ " を仮定してみます．この仮説を帰無仮説といいます．この帰無仮説を仮定すると，確率変数 $T=\dfrac{X-m}{s}$ は標準正規分布 $\mathrm{N}(0,1)$ に従うことになるので，T の実現値が 0 から離れれば離れるほど帰無仮説は疑わしくなります．そこで，確率が 1% 以下になるほど T の実現値が 0 から離れるとき，帰無仮説は成り立たないと一応判断します．帰無仮説の下で，$\mathrm{P}(\,|T|\ge 1.960\,)=0.05$ なので，T の実現値 t について $|t|\ge 1.960$ となる確率は 5% です．それ故，$|t|\ge 1.960$ となるとき，帰無仮説を否定します．このことを " 帰無仮説を棄却する " といいます．$|t|\le 1.960$ となるとき，帰無仮説は否定できません．

このように，統計量 T の実現値 t によって帰無仮説を棄却するかしないかが決まります．このような統計量 T を**検定統計量** (test statistic) といい，棄却するかどうかを考える元の仮説を**帰無仮説** (null hypothesis) といいます．帰無仮説は成り立たないと一応判断することを " 帰無仮説を**棄却する** (reject) " といい，帰無仮説を棄却することになるような検定統計量の実現値の範囲を**棄却域** (rejection region) といいます．つまり次のようにします：

検定統計量の実現値が棄却域に属す ならば 帰無仮説を棄却する；
検定統計量の実現値が棄却域に属さない ならば 帰無仮説を棄却しない．

この例では，検定統計量 T の実現値 t について，帰無仮説を棄却する条件は，$|t|\ge 1.960$ ，つまり，$t\le -1.960$ または $t\ge 1.960$ です．よって棄却域は区間の合併 $(-\infty,-1.960]\cup[1.960,\infty)$ です．この棄却域は，検定統計量 T の実現値 t が棄却域に属す確率が 0.05 になるように定めました：$\mathrm{P}(\,|T|\ge 1.960\,)=0.05$ ．この棄却域を定める基準となる確率の値 0.05 を**有意水準** (significance level) あるいは危険率といいます．有意水準とは，帰無仮説が本当であるにも関わらず帰無仮説を棄却してしまう危険性を表す確率です．有意水準は，多くの場合，$0.05=5\%$ または $0.01=1\%$ に設定します．

このように，与えられた有意水準の下で帰無仮説が棄却できるかどうか判断することを統計的仮説検定あるいは単に**仮説検定** (hypothesis testing) といいます．

仮説検定において，帰無仮説の否定を**対立仮説** (alternative hypothesis) といいます．帰無仮説を棄却するということは，帰無仮説を否定する，つまり対立仮説を肯定することです．帰無仮説を棄却できないとき，" 帰無仮説を採択する (accept) " といいます．帰無仮説を採択するということは，帰無仮説が正しい可能性を無視できないということであり，帰無仮説が正しいと積極的に主張しているのではありません [1]．ですから，帰無仮説を棄却できないときは，帰無仮説が正しいとも誤りであるとも主張できません．

一般的に，仮説検定では以下の手順を踏みます．

(1) 帰無仮説を立てる．
(2) 検定統計量を定める．
(3) 与えられた有意水準から棄却域を定める．
(4) 検定統計量の実現値を求める．
(5) 検定統計量の実現値が棄却域に属すならば帰無仮説を棄却して対立仮説を主張する；属さないならば帰無仮説を採択する．

特に (3) と (4) との順序に注意して下さい．必ず検定統計量の実現値を求める前に棄却域を定めなければなりません．

---

[1] 例えば，刑事裁判において，有罪にする証拠が充分でなければ無罪になります．無罪ということは，無実である可能性を捨て切れないということであり，無実であると積極的に主張している訳ではありません