確率統計 両側検定と片側検定（前回の解答）　2019.12. 3 分

平均に関する仮説検定の方法はおおよそ次のとおりである。

1）仮説を立てる。

　平均に関する検定を行う場合、調べることは平均がある値 α と「同じかどうか」もしくは「それよりも大きい（小さい）」である。もし「同じかどうか」であれば、

帰無仮説 H_0:同じである $(\mu = \mu_0)$
対立仮説 H_1:異なる $(\mu \neq \mu_0)$

を仮説とし、「それよりも大きい（小さい）」であれば

帰無仮説 H_0:同じか小さい $(\mu \leqq \mu_0)$ 《もしくは同じか大きい $(\mu \geqq \mu_0)$》
対立仮説 H_1:大きい $(\mu > \mu_0)$ 《もしくは小さい $(\mu < \mu_0)$》

を仮説とすればよい。つまり前回の解答は「(添加後の燃費) − (添加前の燃費)」なので

帰無仮説 H_0:燃費は同じか悪くなった $(\mu \leqq 0)$ 《実際には燃費は同じでも可 $(\mu = 0)$》
対立仮説 H_1:燃費は良くなった $(\mu > 0)$

2）母集団から標本を抽出する。

　基本は無作為抽出でよいが、調べたい事柄によっては層別抽出などを用いる。

3）帰無仮説を真と仮定したときの、標本の出現確率を計算する。

　実際には標本平均 $\bar{x}$ が α に近いかどうかで判断するために、確率を用いてその基準（棄却域）を作る。母分散が既知か未知かによって、

$$Z = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} \quad \text{もしくは} \quad T = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}}$$

で変換して、両側検定の場合（対立仮説が $\neq$ のとき）

$$\Pr\{|\, Z \text{ or } T\,| > \varepsilon\} = \alpha \text{ (有意水準)}$$

となる ε を求め、片側検定の場合（対立仮説が $<, >$ のとき）

$$\Pr\{(\, Z \text{ or } T\,) > \varepsilon\} = \alpha \text{ (有意水準)} \quad \text{（対立仮説が } \mu > \mu_0 \text{ のとき）}$$
$$\Pr\{(\, Z \text{ or } T\,) < \varepsilon\} = \alpha \text{ (有意水準)} \quad \text{（対立仮説が } \mu < \mu_0 \text{ のとき）}$$

となる ε を求める。もしくはコンピュータ等で直接次の確率を計算する

$$\Pr\{|(\, Z \text{ or } T\,)| > (z \text{ or } t)\} \text{ (両側検定の場合)} \quad \text{（対立仮説が } \mu = \mu_0 \text{ のとき）}$$
$$\Pr\{(\, Z \text{ or } T\,) < (z \text{ or } t)\} \text{ (片側検定の場合)} \quad \text{（対立仮説が } \mu < \mu_0 \text{ のとき）}$$

4）3）の出現確率が極めて低いとき、帰無仮説を棄却して対立仮説が正しいと判定する。
そうでないときは、判断を保留する。
つまり、下記の図の灰色部分に入ったとき、帰無仮説を棄却すればよい。

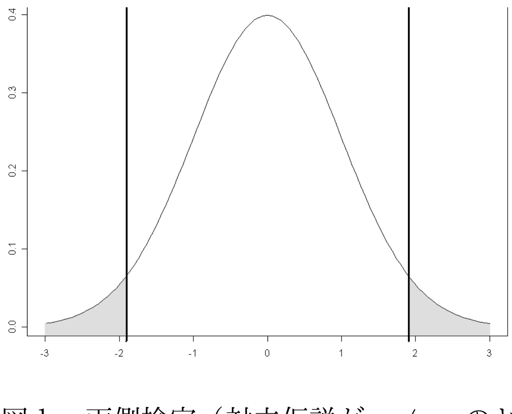

図１．両側検定（対立仮説が $\mu \neq \mu_0$ のとき）

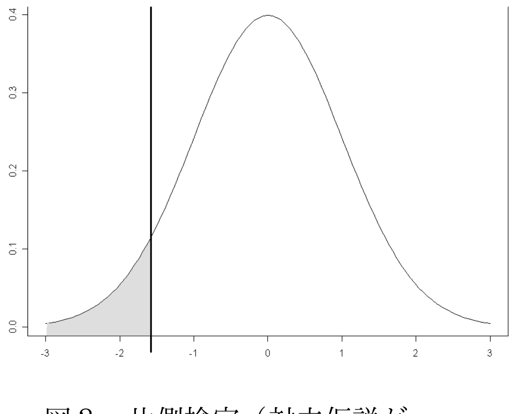

図２．片側検定（対立仮説が $\mu < \mu_0$ のとき）

例．人種による身長の差 (両側検定)

オーストラリア人の新生児１０人を無作為に抽出し、その身長を測定したところ次のようであった。

51.0　45.9　48.8　54.0　53.5　48.0　44.5　46.0　50.3　48.0

日本人の新生児の平均身長が 50.2cm であったとしたとき、日本人とオーストラリア人の新生児の平均身長が異なるかどうかを調べる。ただし、身長の分布は正規分布とみなしてよいとする。

これを前ページの順番に沿って考えると。

１）仮説を立てる

この問題では、日本人とオーストラリア人の新生児の平均身長に差があるかとどうかを調べている。一般的にこのような問題では差がある方に興味があることが多いので、帰無仮説と対立仮説は次のようになる。

帰無仮説 H_0：平均身長に差は無い
対立仮説 H_1：平均身長に差がある

このように仮説をおけば、帰無仮説か対立仮説のいずれかが正しいことになる。つまり帰無仮説が棄却された場合、対立仮説が正しいと判断できる。これを数式で表すと、オーストラリア人の新生児の平均身長を μ として

帰無仮説 H_0：$\mu = 50.2$
対立仮説 H_1：$\mu \neq 50.2$

となる。

２）母集団から標本を抽出する。

既に無作為抽出されているので省略。

本来であれば、仮説を立ててから標本を抽出するのが本筋であるが、無作為抽出の場合、仮説が抽出に与える影響がないので、特に順番が入れ替わっても問題は無い。しかし、仮説によっては層別抽出などの必要性が生じることがあるため、仮説を立ててから標本を抽出するのが正しい。

３）帰無仮説を真と仮定したときの、標本の出現確率を計算する。

帰無仮説を真と仮定するので、オーストラリア人の身長は正規分布 $N(50.2, \sigma^2)$ に従っていると仮定する。その仮定の下で、上記１０人の身長どの程度の確率で出現するかを計算する必要があるが、標本ごとに個別に確率を計算するのではなく、１０人まとめて平均値 $\bar{x}$ の出現確率を用いて判断することになる。つまり、標本平均の値 $\bar{x} = 49.0$ の出現確率が問題となる。推定のところでも出てきたように、標本平均の分布は元の分布と母平均が等しく、母分散を標本数で割った値になることが分かっているので、この場合 $\bar{x}$ は $N(50.2, \frac{\sigma^2}{10})$ に従うときの、$P(\bar{X} < 49.0)$ の確率を求めればよい。

４）３）の出現確率が極めて低いとき、帰無仮説を棄却して対立仮説が正しいと判定する。そうでないときは、判断を保留する。

３）で求めた出現確率が極めて低いとあるが、その境はどこにあるのであろうか。この答えは「問題によって異なる」としか答えようが無い。一般的には５％（２０回に１回）もしくは１％（１００回に１回）程度の出現確率のときに、極めて低い確率と見るが、人の命がかかっている場合等は０．１％（千回に１回）や０．０１％(一万回に１回) の出現確率でやっと極めて低いとみなす場合がある。確率は０でない限り、起こる可能性がある。またこの値は、実際に帰無仮説が真の場合に帰無仮説を棄却する第１種の過誤の確率と等しくなる。この確率の値は事前に有意水準と呼ばれる値で決めておき、出現確率が有意水準以下の確率になったとき、帰無仮説を棄却する。

今回の例では特に有意水準の指定は無いが、ここでは有意水準５％として話を進める。３）の最後に記載したように $P(\bar{X} < 49.0)$ の確率を求めることが可能であれば、その値と有意水準を比べて帰無仮説を棄却するかどうか判断すればよい。しかしながら今回の例の場合のように、母分散が未知の場合は t 分布を使うことになるので、一般的に仮説検定では次のような形で、帰無仮説を棄却するかどうか判断する。

a) $Z = \dfrac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$ もしくは $T = \dfrac{\bar{X}-\mu}{\frac{S}{\sqrt{n}}}$ で変換する。推定問題と異なり、検定の場合すべての変数 $\bar{x}, \mu, \sigma^2(\text{ or } s^2), n$ の値がわかっているので、実際に値が求まる。

b) 有意水準を α としたとき、
　z であれば標準正規分布における $P(z < -z_\alpha$ もしくは $z_\alpha < z) = \alpha$
　t であれば自由度 $n-1$ の t 分布における $P(t < -t_\alpha$ もしくは $t_\alpha < t) = \alpha$
となる t_α の値を求める。

c) $|z|$ もしくは $|t|$ の値と t_α の値を比較して
　$|z| > z_\alpha$ もしくは $|t| > t_\alpha$ であれば、帰無仮説を棄却して、
　$|z| \leq z_\alpha$ もしくは $|t| \leq t_\alpha$ であれば判断を保留する。

実際に計算をしてみると、母分散が未知なので、標本分散 $s^2 = (3.194)^2$ を用いて変換すると t の値は

$$t = \frac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}} = \frac{49.0-50.2}{\frac{3.194}{\sqrt{10}}} = -1.18808$$

となる。この t は自由度 $10-1=9$ の t 分布に従うので、$P(t < -t_\alpha$ もしくは $t_\alpha < t) = 0.05$ となる t_α の値は 2.262 である。よって、$|t|$ が 2.262 よりも大きいときは標本の出現確率が５％以下になるために、帰無仮説を棄却するが、この例の場合は $|t| \leq 2.262$ なので帰無仮説を棄却することができない。つまり判断を保留する。

この例の場合、棄却する箇所が分布の両側にあるので、両側検定と呼ばれる。両側検定は基本的に $H_0 : \mu = \mu_0$ と $H_1 : \mu \neq \mu_0$ のように、単純に異なるかどうかを検定したい場合に用いられる。

2019年度 確率統計　小テスト解答用紙　2019.12.10

・検定　1

　ある会社のＡ工場で作られた電球の寿命を調べたら下記の通りであった。

84.3, 89.4, 84.5, 87.3, 89.2, 84.2, 82.4

Ａ工場とは別のＢ工場で作られた電球の寿命の平均が $\mu = 83.1$ であるとき、両工場で作られる電球の寿命に差があるかどうか有意水準５％で検定せよ。但し、電球の寿命は正規分布に従うと仮定する。

2019年度神奈川工科大学 確率統計 演習問題	学科	学年	組	学　籍　番　号	氏　　　　名	

提出先：K3-3309号室前　17番のボックス　　提出期限：１２月１１日（水）１７時頃まで