

$$-1.96\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} < (\bar{x}_1 - \bar{x}_2) - (m_1 - m_2) < 1.96\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

これを、 $(m_1 - m_2)$  について解いて、

$$(\bar{x}_1 - \bar{x}_2) - 1.96\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} < m_1 - m_2 < (\bar{x}_1 - \bar{x}_2) + 1.96\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \dots (イ)$$

もとの二つの母集団  $A_1, A_2$  が正規分布をなし、 $\sigma_1, \sigma_2$  が既知の場合は、この不等式を用いて、信頼度95%の母平均の差の区間推定を行えばよいわけですが、一般には母集団の分布が不明、 $\sigma_1, \sigma_2$  も未知な場合がふつうです。それで、このような場合には、 $n_1, n_2$  をともに大、目安として100以上にとりますと、上の不等式(イ)の $\sigma_1, \sigma_2$  の代わりに、標本標準偏差  $s_1, s_2$  を代用しても、そう困るほどの差がないことが、経験的にわかっていますので、

$$\bar{x}_1 - \bar{x}_2 - 1.96\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} < m_1 - m_2 < \bar{x}_1 - \bar{x}_2 + 1.96\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \dots (ロ)$$

とし、二つの母平均の差  $m_1 - m_2$  の信頼度95%の信頼区間として、

$$(\bar{x}_1 - \bar{x}_2 - 1.96\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}, \bar{x}_1 - \bar{x}_2 + 1.96\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}})$$

を用いることができます。

ここで、すでにお気づきになっていると思いますが、次のことを注意しておきます。それは、95%の確率で信用できる、ということと、信用できない確率が5%ある、ということとは、うらはらの関係で、その意味内容は全く同じということです。

したがって、信頼度95%と、危険率5%とは、どちらを表だてて使うかによって使い分けられているだけで、その意味内容は全く同じことなのです。すなわち、区間推定の場合には、信頼度を表だてて用い、検定の場合には、危険率(有意水準)を表だてて用いる習慣になっています。

さて、話をもとにもどしまして、[4] による母平均の差の検定で  $|z| \geq 1.96$  となり、危険率5%で有意差あり、となったとします。そうしますと、

$$|z| = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \geq 1.96 \quad \text{より}$$