2標本の平均値の検定に用いる統計量t

次に，研究や技術開発において有意性検定を行なう状況を考える．Fisher (1926) では，施肥法の違いにより 10%の増収があった場合を例に有意性を解説している．ここでは具体的な数値で説明する．水稲の収量に関して，堆肥を用いた新しい施肥法（処理 A）と従来の施肥法（処理 B）の 2 つの処理を考える．それぞれ 5 反復の完全無作為化法で実験を行ない，標本平均と標準偏差の推定値

$$\bar{x}_A = 550, \quad \bar{x}_B = 500, \quad \hat{\sigma} = 25$$

が得られたとする（単位は kg/10a，以下単位は省略）．増収効果 $\mu_A - \mu_B$ の推定値は

$$\bar{x}_A - \bar{x}_B = 50 \quad (\bar{x}_B \text{ に対して 10\%の増収})$$

である．我々は推定値が変動することを知っている．したがって

- 従来の施肥法 B のもと（すなわち帰無仮説 $H_0: \mu_A = \mu_B$ のもと）でも，これくらいの収量（あるいはそれ以上の収量）が生じることがあるのではないか

という疑問を抱くことは自然なことである．この疑問に対して，伝統的な t 検定では

$$\boxed{|t| = \frac{|\bar{x}_A - \bar{x}_B|}{\hat{\sigma}\sqrt{2/n}} = 3.16 \quad (p = 0.013)}$$

が得られる（ここでは減収の場合も含めて両側検定を考えている）．

2 つの処理を取り上げて実験を行なったときに我々が知りたいことは，「どれぐらいの違いがあるのか」，また「それは本当に処理の違いによる効果なのか」ということである．最初の「どれぐらいの違いがあるのか」という問いに対しては，$\bar{x}_A - \bar{x}_B = 50$ が示している．一方，t 統計量（あるいは，それと 1 対 1 の関係にある p 値）が示していることは

$$\boxed{\frac{\text{効果の推定値}}{\text{誤差の推定値}}}$$

すなわち SN 比 (signal-to-noize ratio) である．この t 統計量（あるいは p 値）を計算することによって，観測された効果の推定値が帰無仮説のもとでも得られるものなのかどうかを判断することができる．

確かに，この実験に対して

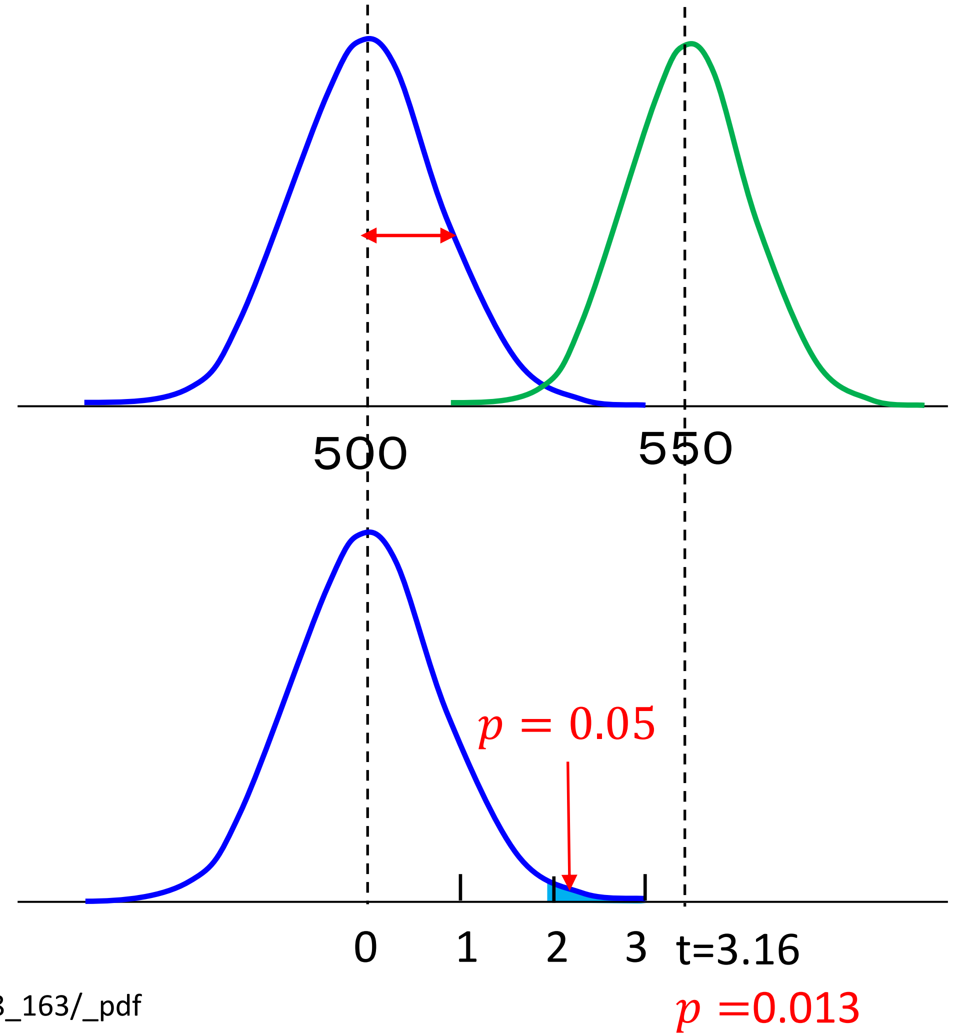
$$|t| = 3.16 \quad (p = 0.013)$$

のように t 統計量のみ（あるいは p 値のみ）を示すことは意味がない．効果の推定値とともに

$$\bar{x}_A - \bar{x}_B = 50 \quad (\text{両側 } p = 0.013)$$

と示すことによって，p 値は有効な情報となる．

$$t = \frac{\bar{x}_A - \bar{x}_B}{\sigma\sqrt{\dfrac{1}{n_A} + \dfrac{1}{n_B}}} \quad \begin{matrix} \leftarrow \text{効果の推定値} \\ \leftarrow \text{誤差の推定値} \end{matrix} \Bigg\} \text{SN比}$$



出典：https://www.jstage.jst.go.jp/article/jjb/38/2/38_163/_pdf