仮説検定

１．検定の手続き

(i)　パラメータに関する互いに排反な 2 組の仮説を設定する。一方を帰無仮説といい、正しいものと仮定され、検定されるべき仮説である。他方は、帰無仮説の一部あるいは全部を否定したもので、対立仮説という。

$H_0 : \mu = \mu_0 \quad \text{vs} \quad H_1 : \mu > \mu_0$

(ii)　適切な検定統計量を選択して、帰無仮説の元での分布（帰無分布）を求め、検定統計値を計算する。

(iii)　有意水準αを設定する。

(iv)　対立仮説の形に対応して、帰無仮説の受容域と棄却域を決める。その際、帰無仮説のもとで検定統計量が棄却域に入る確率を有意水準と等しくなるように、受容域と棄却域の境界点（臨界値、有意点）を決める。

(v)　実際のデータから検定統計値を計算して、その値が有意点を超えるときには帰無仮説を棄却し、さもなければ受容する。

検定においては、(i)にあるように 2 つの仮説を用意する。例えば、ある新薬を開発した製薬会社が、その新薬が旧薬よりも優れているかどうかを調べるものとしよう。ここで、新薬、旧薬の効能の水準を、それぞれμ, μ_0とする。旧薬はすでに市場に出回って使われているので、μ_0は既知である。しかし、新薬は臨床実験が済んだ段階であるので、μは未知であるとする。しかし、新薬が旧薬よりも劣ることはないものとしよう。このときの検定問題は、

$$H_0 : \mu = \mu_0 \quad \text{vs} \quad H_1 : \mu > \mu_0$$

と表すことができる。ここで、H_0は帰無仮説、H_1は対立仮説である。帰無仮説は、新薬の効能は従来通りということ、対立仮説は、新薬の方が優れているということを意味している。検定における仮説の受容、棄却という場合の仮説は、帰無仮説を意味する。そして、いずれを結論するにしても、帰無仮説が正しいという想定のもとで推論が展開される。したがって、帰無仮説は推論を展開しやすい形に設定するべきである。上の例で、帰無仮説を新薬の方が優れているとした場合、どの程度優れているかを細かく設定しない限り、推論を展開することは出来ない。条件が明確な上記のような帰無仮説の方が、はるかに適切であることが了解されよう。

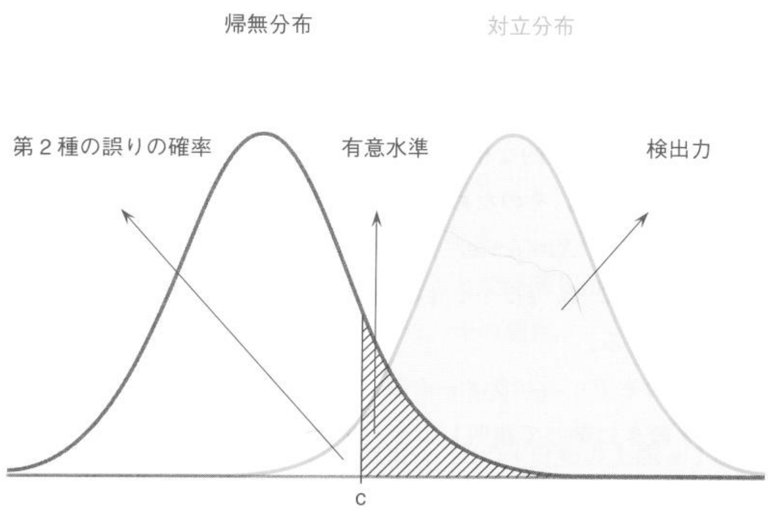
以上の理由により、論理的には、帰無仮説の受容が帰無仮説の正しさを証明したことにはならない。帰無仮説とは、まさに無に帰したい仮説であり、帰無仮説が棄却されたときはじめて、帰無仮説の正しさではなく、間違いを積極的に主張できることになる。

２．検定における２種の誤り：

		帰無仮説の状態	
		真	偽
帰無仮説の採否	受容	○	第２種の誤り
	棄却	第１種の誤り	○

帰無仮説を受容するにしても棄却するにしても、それぞれの結論には常に誤りが付きまとっている。なぜなら帰無仮説を棄却する際には第一種の誤り（帰無仮説が真であるにもかかわらず棄却する誤り）が、そして受容する際には第 2 種の誤り（帰無仮説が偽であるにもかかわらず受容する誤り）が起こりうる。これら二つの誤りはトレード・オフの関係にある。それで、検定においては、第一種の誤りの確率を一定の低い水準にコントロールしておいて、第 2 種の誤りの確率をなるべく小さくするように受容域と棄却域を決める方式がよい。
有意水準αとは第一種の誤りの確率を表す数値である。通常 0.01、0.05、0.10 などの区切りがよい数値が使われる。有意水準を高めれば、帰無仮説は棄却されやすくなる。しかし、その分だけ、従来通り（$H_0 : \mu = \mu_0$）可も知れない可能性を棄てる危険を高めることにつながる。この意味で有意水準を危険率とも呼ばれる。
下の図で左側は、帰無仮説での統計量分布である。右側は、対立仮説での統計量分布である。帰無仮説を棄却するかどうかの基準になる値 c は臨界値（Critical　Value）と呼ばれる。 c 以下の領域は受容域、 c を超える領域は棄却域。

第１種の誤りの確率で有意水準αでもある。$\alpha = P(H_0$を棄却｜$H_0)$
第２種の誤りの確率は、　$\beta = P(H_1$を棄却｜$H_a)$
検出力＝１－β（第２種の誤りの確率）＝対立仮説が真の時帰無仮説を棄却する確率



３．平均検定の検定３つのパタン

正規母集団の平均μに関する仮説検定で、単純確率モデル

$X_i = \mu + \varepsilon_i \quad (i = 1, \ldots, n)$ 、$\varepsilon_i \sim N(0, \sigma^2)$

(i)　　両側検定

$H_0 : \mu = \mu_0 \quad \text{vs} \quad H_1 : \mu \neq \mu_0$

(ii)　　片側検定

$H_0 : \mu = \mu_0 \quad \text{vs} \quad H_1 : \mu > \mu_0$　右肩側検定

or

$H_0 : \mu = \mu_0 \quad \text{vs} \quad H_1 : \mu < \mu_0$　左肩側検定

４．平均検定の手順

(i)帰無仮説と対立仮説の設定

$H_0 : \mu = \mu_0 \quad \text{vs} \quad H_1 : \mu \neq \mu_0$　両側検定

or

$H_0 : \mu = \mu_0 \quad \text{vs} \quad H_1 : \mu > \mu_0$　右肩側検定

or

$H_0 : \mu = \mu_0 \quad \text{vs} \quad H_1 : \mu < \mu_0$　左肩側検定

(ii)検定統計量の選択:　母分散(σ^2)の既知または未知によって二つの検定統計量を選択
平均μに関する検定であるから、μの推定量$(\bar{X})$を使えばよいと考えられる。$\bar{X}$の帰無分布は、$N(\mu_0, \frac{\sigma^2}{n})$であり、標準化により、検定統計量は、

母分散の既知：標準正規分布を使う：$\dfrac{\sqrt{n}\left(\bar{X} - \mu_0\right)}{\sigma}$ 〜 $N(0,1)$

母分散の未知：ｔ分布を使う：$\dfrac{\sqrt{n}\left(\bar{X} - \mu_0\right)}{\hat{\sigma}}$ 〜 t (n-1)

$$\hat{\sigma} = \sqrt{\frac{\sum (X_i - \bar{X})^2}{n-1}}$$

(iii)　　有意水準αの設定、通常$\alpha = 0.05, \quad \alpha = 0.01$

(iv)　有意水準αにしたがって受容域と棄却域の設定。

$H_0: \mu = \mu_0$　vs　$H_1: \mu \neq \mu_0$　両側検定

対立仮説の形から、帰無仮説は検定統計量の絶対値が大きくなるときに棄却される。すなわち。$\dfrac{\sqrt{n}\left|\overline{X} - \mu_0\right|}{\sigma} > c$ となるときに帰無仮説を棄却する。　この有意点（または臨界値）は、検定統計量が棄却域に入る確率を帰無分布の元で評価したときに、有意水準と等しくさせるような点である。

$P(\dfrac{\sqrt{n}\left|\overline{X} - \mu_0\right|}{\sigma} > c \mid H_0) = \alpha$ を満たす値 c である。　c は、標準正規分布の上側$\dfrac{\alpha}{2}$点、$z_{\frac{\alpha}{2}}$

となる。受容域：$\dfrac{\sqrt{n}\left|\overline{X} - \mu_0\right|}{\sigma} \leq z_{\frac{\alpha}{2}}$　　棄却域：$\dfrac{\sqrt{n}\left|\overline{X} - \mu_0\right|}{\sigma} > z_{\frac{\alpha}{2}}$

(V) 決定

検定統計量を計算し、有意水準が示す棄却域にはいるのか受容域にはいるのかを確認し、帰無仮説を棄却もしくわ採択する。

５．平均検定、例１：

新生児の体重の分布は正規分布で近似できる。その平均体重はほぼ 3.1kg であると知られている。標準偏差は、0.2kg としよう。ある新生児のグループについて体重が 3.1kg より低いのではないか、という疑問が生じた。新生児 25 人について観測値を求め、その平均は 3.0kg で分散は 0.05 であった。この疑問について統計的に検定し、推論しよう。

１）帰無・対立仮説設定

$H_0: \mu = 3.1$、$H_a: \mu < 3.1$

２）帰無仮説元での検定統計量・検定統計値：母分散が既知と未知

母分散既知の場合、標本統計値：$Z_{H_0} = \dfrac{\overline{X} - \mu_{H_0}}{\sigma/\sqrt{n}} \dfrac{3.0 - 3.1}{0.2/\sqrt{25}} = \dfrac{-0.1}{0.04} = -2.5$

母分散が未知の場合、標本統計値：$t_{H_0} = \dfrac{\overline{X} - \mu_{H_0}}{s/\sqrt{n}} = \dfrac{3.0 - 3.1}{0.224/\sqrt{25}} = -2.23$

３）有意水準５％での臨界値：$P(Z_0 \le -1.65 \mid H_0) = 0.05$

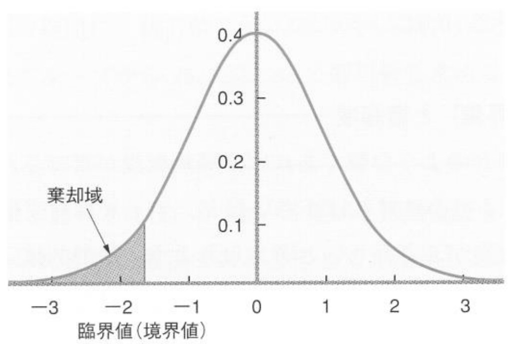
既知の場合：標本正規分布表から：$P(Z_0 \le -1.65 \mid H_0) = 0.05$

$Z_{0.05} = 1.65 \rightarrow -Z_{0.05} = -1.65$

未知の場合：ｔ分布表から：$P(t_0 \le -1.711 \mid H_0) = 0.05$

$t_{0.05}(24) = 1.711 \rightarrow -t_{0.05}(24) = -1.711$

図6.1　帰無仮説下での Z_0 の分布



４）決定：採択か棄却か

$$Z_{H_0} = \frac{3.0 - 3.1}{0.2/\sqrt{25}} = -2.5 < -Z(0.05) = -1.65$$ ：帰無仮説棄却

$$t_{H_0} = \frac{\overline{X} - 3.1}{s/\sqrt{25}} = -2.23 < -t_{0.05}(24) = -1.711$$ ：帰無仮説棄却

５）解釈：有意な結果であり、新生児体重の平均が 3.1kg より小さくなったとの仮説が認められた。

例２：

1999 年度の学校保育統計調査報告書によると、滋賀県の１７歳児の平均身長は、男子 171.7cm、女子 158.6 であった。男女とも観測個数ｎは 450 であった。なお、1949 年の 17 歳児の平均身長は、男子 161.2cm、女子 152.2cm である。それで、平均身長が、この 50 年間伸びたといえるのか。標準偏差は、男子は 5.74cm、女子は 5.22cm であるとする。母分布は未知であるが、身長の平均分布は標本が大きくなるつれ、正規分布にしたがう（中心極限定理）。

１）帰無・対立仮説設定

$H_0 : \mu = 161.2$、$H_a : \mu > 161.2$

２）帰無仮説元での検定統計量・検定統計値：母分散が既知であるから

母分散が既知の場合、標本統計値：$Z_{H_0} = \dfrac{\overline{X} - \mu_{H_0}}{\sigma/\sqrt{n}} = \dfrac{171.7 - 161.2}{5.74/\sqrt{450}} = 38.8$

３）有意水準１％での臨界値：標準正規分布表から $P(Z_0 \geq 2.3 | H_0) = 0.01$

$Z_{0.01} = 2.3$

４）決定：$Z_{H_0} = \dfrac{171.7 - 161.2}{5.74/\sqrt{450}} = 38.8 > 2.3 = Z_{0.01}$ ：帰無仮説棄却

５）有意な結果であり、平均身長は伸びたといえる。

例３：Page187

ある工場で生産された電球の平均寿命は 1200 時間である。この工場で新しい生産工程が稼動し初めてたが、この新工程から生産された電球を 100 個無作為抽出し、電球の平均寿命 μ が伸びたかどうか検定する。有意水準は１％にする。得られた標本からの標本平均は 1230 時間、標準偏差は標本観測値より 120 と推定される。母集団分布は分からないが、ｎが 100 であるので、標本平均の分布は正規分布によって近似できよう。

１）仮説設定：$H_0 : \mu = 1200$、$H_a : \mu > 1200$

２）検定統計量と検定統計値：母分散未知であるから

$$t_{H_0} = \frac{\overline{X} - \mu_{H_0}}{s/\sqrt{n}} = \frac{1230 - 1200}{120/\sqrt{100}} = 2.5$$

３）有意水準１％での臨界値：ｎが 100 個でｔ分布は標準正規分布に近似できる。

標準正規分布表から $P(Z_0 \geq 2.3 | H_0) = 0.01$ ：$Z_{0.01} = 2.3$

４）決定

$$t_{H_0} = \frac{1230 - 1200}{120/\sqrt{100}} = 2.5 > 2.3$$

５）解釈：有意な結果であり、平均寿命は新しい工程の導入で伸びたと言える。

６．平均差の検定

二つの異なる正規母集団があり、各々の母集団において平均値が μ_i、標準偏差は σ_i、i=1,2.
このような状況で平均値 μ_1、μ_2 が等しいかどうか検定する。
中心統計量は標本平均の差：$\overline{X} = \overline{X}_1 - \overline{X}_2$。$E(\overline{X}) = E(\overline{X}_1) - E(\overline{X}_2) = \mu_1 - \mu_2$

I)母分散が既知での母平均差の検定

$$\overline{X}=\overline{X}_1-\overline{X}_2 \sim N(\mu_1-\mu_2、\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2})、$$

$$V(\overline{X})=V(\overline{X}_1)+V(\overline{X}_2)=\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}$$ ：母分散が既知

１）仮説設定：$H_0:\mu_1=\mu_2$、$H_a:\mu_1\neq\mu_2$

２）定統計量：$Z_0=\dfrac{\overline{X}}{\sqrt{V(\overline{\overline{X}})}}=\dfrac{\overline{X}_1-\overline{X}_2-(\mu_1-\mu_2)_{H_0}}{\sqrt{V(\overline{\overline{X}})}}=\dfrac{\overline{X}_1-\overline{X}_2}{\dfrac{\sigma^2{}_1}{n_1}+\dfrac{\sigma^2{}_2}{n_2}}$ ～N(0,1)

３）有意水準α％での棄却域：$H_a:\mu_1\neq\mu_2$のとき　$Z_0=\dfrac{\left|\overline{X}\right|}{\sqrt{V(\overline{\overline{X}})}}>Z_{\alpha/2}^{(n-1)}$

例４：平均差の検定、例 6.4（191Page）

1998 年度の学校保育統計調査報告書によると、１７歳の男子平均身長が一番高い県は秋田県で 172.1、標準偏差は 5.94 であった。平均身長が一番低い県は沖縄県で 168.9、標準偏差は 5.90 であった。観測個数は各々430 とする。秋田県からの分散（σ_1^2）は３５３で沖縄県からの分散（σ_2^2）は 34.8 である。ここで平均の最高値と最低値の間に有意な差があるかどうか有意水準１％で検定せよ。

（１）　仮説設定：$H_0:\mu_1=\mu_2$、$H_a:\mu_1>\mu_2$

（２）　検定定統計量と検定統計値：母分散既知

$$V(\overline{X})=V(\overline{X}_1)+V(\overline{X}_2)=\frac{35.3}{430}+\frac{34.8}{430}$$

母分散既知：標準正規分布適用

$$Z_0=\frac{\overline{X}}{\sqrt{V(\overline{X})}}=\frac{172.1-168.9}{\sqrt{\frac{1}{430}(35.3+34.8)}}=7.9$$

（３）　決定：$Z_0=\dfrac{\overline{X}}{\sqrt{V(\overline{X})}}=7.9>Z(0.01)=2.3$

（４） 解釈：平均の差は有意であった。

II) 母分散が未知であるが同じである時の母平均差の検定

$\overline{X} = \overline{X}_1 - \overline{X}_2 \sim N(\mu_1 - \mu_2$、$S^2(\frac{1}{n_1} + \frac{1}{n_2}))$、母分散が未知であるが同じ

$\hat{V}(\overline{X}) = S^2(\frac{1}{n_1} + \frac{1}{n_2})$、$S^2 = \frac{1}{n_1 + n_2 - 2}\{\sum_{i=1}^{n_1}(X_{1i} - \overline{X}_1)^2 + \sum_{i=1}^{n_2}(X_{2i} - \overline{X}_2)^2\}$

$S^2 = \frac{1}{n_1 + n_2 - 2}\left((n_1 - 1)S^2{}_1 + (n_2 - 1)S^2{}_2\right)$、$S_i^2 = \frac{\sum_{i=1}^{n_i}(X_{1i} - \overline{X})^2}{n_i - 1}$、標本分散

１）仮説設定：$H_0 : \mu_1 = \mu_2$、$H_a : \mu_1 \neq \mu_2$

２）検定統計量と検定統計値：母分散未知ｔ分布適用

$$t_0 = \frac{\overline{X}}{\sqrt{\hat{V}(\overline{X})}} = \frac{\overline{X}_1 - \overline{X}_2 - (\mu_1 - \mu_2)_{H_0}}{\sqrt{\hat{V}(\overline{X})}} = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{S^2(\frac{1}{n_1} + \frac{1}{n_2})}} \sim t(n_1 + n_2 - 2)$$

３）有意水準α％での棄却域：$H_a : \mu_1 \neq \mu_2$のとき　$t_0 = \frac{|\overline{X}|}{\sqrt{\hat{V}(\overline{X})}} > t_{\alpha/2}^{(n_1+n_2-1)}$

例５）母平均差の検定：母分散が未知であるが同じ

A,B の 2 たつの大学から男子学生をランダムにそれぞれ 17 名と 14 名を選んで身長を測定し、各種統計量を計算したところ次のようになった。

大学	サンプル数	平均	標本分散（$S^2{}_i$）
A	17	168.8	26.32
B	14	165.0	25.87

平均身長の差はあるといえるのか。有意水準 5%検定せよ。

１）仮説設定：$H_0 : \mu_1 = \mu_2$、$H_a : \mu_1 \neq \mu_2$

２）検定統計量と検定統計値：母分散未知ｔ分布適用

$$t_0 = \frac{\overline{X}}{\sqrt{\hat{V}(\overline{X})}} = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{S^2(\frac{1}{n_1} + \frac{1}{n_2})}} = \frac{168.8 - 165.0}{\sqrt{26.118(\frac{1}{17} + \frac{1}{14})}} = 2.060 > 2.045 \sim t_{0.025}(29)$$

$$S^2 = \frac{1}{n_1 + n_2 - 2}\left((n_1 - 1)S^2{}_1 + (n_2 - 1)S^2{}_2\right) = \frac{1}{17 + 14 - 2}(16 \times 26.32 + 13 \times 25.87) = 26.118$$

３）有意水準５％で帰無仮説棄却。身長の差は有意である。