統計的仮説検定

未知母数を具体的な数値と比較し差異があるか調べたい.

例図 (1) [メーカーの主張] 平均値 $\theta=250$

(2) [消費者の主張] 平均値 $\theta<250$

(3) メーカーの主張を仮定し. 内容量は $N(250,(3.2)^2)$ に従うとする.

すると. 大きさ10の標本平均 $\overline{X}$ は $N(250,\frac{(3.2)^2}{10})$ に従う.

ここで $Z=\frac{\sqrt{10}}{3.2}(\overline{X}-250)$ は $N(0.1)$ に従う.

(4) 小さい値 0.05 に対し.

$P(\overline{X}<\xi)=0.05$ となる ξ を求めると.

$P(\overline{X}<\xi)=P(Z<\frac{\sqrt{10}}{3.2}(\xi-250))=0.05$ より

$P(0<Z<-\frac{\sqrt{10}}{3.2}(\xi-250))=0.05$ となり.

$-\frac{\sqrt{10}}{3.2}(\xi-250)=1.6449$, $\xi\fallingdotseq248.3$ となる.

(5) つまり $\overline{X}=248.2$ は確率的におこりにくい (5%)

$\overline{X}<248.3$ に入っている

(6) 消費者の主張が正しく思える.

<u>母数の検定の一般形式</u>

未知の母数 θ と. 具体的な数 θ_0 の大小関係を考えたとき.

主張されることがらを 統計的仮説 あるいは 仮説 という具体的には

(a) $\theta>\theta_0$ (b) $\theta<\theta_0$ (c) $\theta\neq\theta_0$ (d) $\theta=\theta_0$

の4つである．これらの真偽を標本の結果から判断することを．

統計的仮説検定．あるいは検定という．

検定の方法

(1) とりあえず．主張(d) $\theta=\theta_0$ を認める．

これを帰無仮説といい．H_0 で表す（例では $H_0 : \theta=250$）

(2) これに対し．(1)と対立する仮説を(a)〜(c)から1つ選ぶ

これを対立仮説といい．H_1 で表す（例では $H_1 : \theta<250$）

また(a),(b),(c)をそれぞれ右側，左側，両側検定という．

(3) θ を予想できる統計量 Θ を決める．これを検定統計量という

（例では Θ は $\bar{X}$ または Z）

(4) 小さな値 α を決める． 0.1, 0.05, 0.01 などが一般的

この α を有意水準または危険率という．この α をもとに．H_1 に応じて．

例えば (b) $\theta<\theta_0$ の場合なら．

$P(\Theta<\theta(\alpha))=\alpha$ をみたす $\theta(\alpha)$ を求める．

ここで．領域 $\Theta<\theta(\alpha)$ を H_1 に対する α の棄却域という．

（例では $\alpha=0.05$，$\theta(\alpha)=248.3$）

(5) Θ の実現値を求める （例では $\Theta=248.2$）

(6) Θ が棄却域にあれば．H_0 は棄却され H_1 が採択される．

このとき．α 有意水準では Θ と θ_0 に有意差があるという．（Θ は θ_0 より有意に小さい）

逆に Θ が棄却域になければ．H_0 が採択され．Θ と θ_0 に有意差はないという．

例15　(1) $H_0 : \theta = 5mm$　　(θは母平均)

(2) $H_1 : \theta < 5mm$

(3) $\overline{X}$を標本平均 とすると. $\overline{X}$は $N(\theta, \frac{(0.009)^2}{100})$ に従う. ここで

$Z = \frac{10}{0.009} \cdot (\overline{X} - \theta) = \frac{10^4}{9}(\overline{X} - \theta)$　は　$N(0.1)$ に従う.

このZを検定統計量とする

(4) $P(Z < \theta(0.05)) = 0.05$ より. $\theta(0.05) = -1.6449$ となり.

棄却域は $Z < -1.6449$ である.

(5) Zの実現値は.

$Z = \frac{10^4}{9}(4.998 - 5) \fallingdotseq -2.22$ である

(6) Zの実現値は棄却域にあるので、H_0は有意水準αで棄却される.

⟶ つまり製造方法を変えた方がよい.

問14 16 17

14. (4)で 0.05を0.01 とすればよいので.

$P(\frac{\sqrt{10}}{3.2}(\xi - 250) < Z < 0) = 0.49$ より

$-\frac{\sqrt{10}}{3.2}(\xi - 250) = 2.3263$　∴ $\xi = 247.6$ となる.

(メーカーの主張)

$\overline{X} = 248.2$ は棄却域　$\overline{X} < 247.6$ に入らないので. H_0が採択される

16 (1) $H_0 : \theta = 12.6$

(2) $H_1 : \theta > 12.6$

(3) $\overline{X}$を 標本平均 とすると. $\overline{X}$は $N(\theta, \frac{(1.8)^2}{20})$ に従う.

ここで $Z = \frac{\sqrt{20}}{1.8}(\overline{X} - \theta)$ は $N(0.1)$ に従う.

この Z を検定統計量とする.

(4) $P(Z > \theta(0.05)) = 0.05$ より $\theta(0.05) = 1.6449$. となり

棄却域は $Z > 1.6449$.

(5) Z の実現値は $Z = \frac{\sqrt{20}}{1.8}(13.2 - 12.6) \fallingdotseq 1.49$.

(6) Z は棄却域にないので. H_0 が採択される.

17 (1) $H_0 : \theta = 49.4$ (2) $H_1 : \theta < 49.4$

(3) $\overline{X}$ を標本平均とすると. $\overline{X}$ は $N(\theta, \frac{(6.91)^2}{200})$ に従い.

$Z = \frac{10\sqrt{2}}{6.91}(\overline{X} - \theta)$ は $N(0.1)$ に従う.

この Z を検定統計量とする

(4) 16 と同様に棄却域は $Z < -1.6449$.

(5) 実現値は $Z = \frac{10\sqrt{2}}{6.91}(48.2 - 49.4) \fallingdotseq -2.46$.

(6) H_1 が採択され. 下回るといってよい.