期待値の検定（統計学の復習）

劉慶豊[1]

小樽商科大学

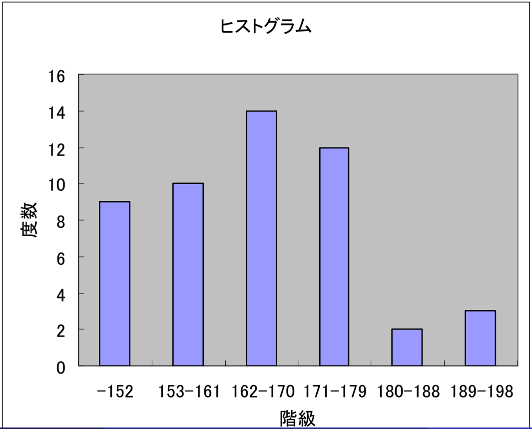
October 2, 2017

[1]E-mail:qliu@res.otaru-uc.ac.jp, URL:http://www.otaru-uc.ac.jp/~qliu/

ヒストグラムと確率密度関数（Probability Density Funtion PDF）

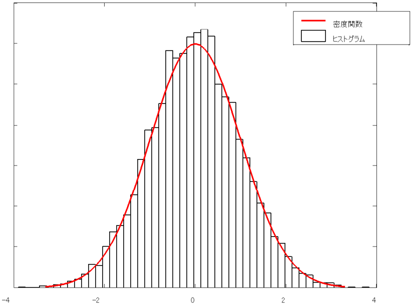
度数 各階級に入っているデータの数．相対度数：度数/全体のデータ数。

各階級の度数を棒グラフにしたものをヒストグラムという。



密度関数

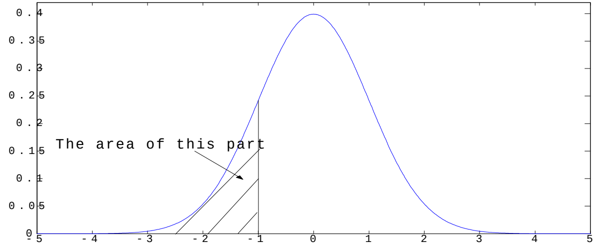
正規分布を例 標準化した身長のデータのヒストグラム（相対度数で描いたもの）、その上に標準正規分布の密度関数を重ね合わせた。正規分布の密度関数のグラフは鐘または富士山の形をしている。



正規分布

密度関数に関して

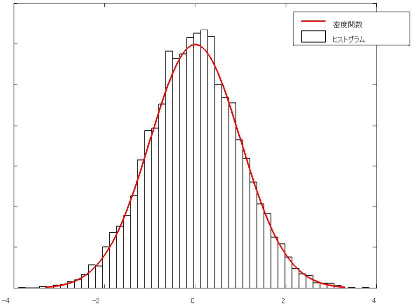
- 密度関数は面積で確率を表現する。
- 密度関数の曲線と X 軸で囲まれた図形の全体の面積は 1 となる。
- 確率変数がある値以下になる確率はその値より左側の密度関数の曲線の下の面積に対応する。
- 確率変数が二つの値の間に入る確率はその二つの値の間の密度関数の曲線の下の面積に対応する。



-1 以下の値になる確率

正規分布の密度関数 期待値がμで分散がσ^2の正規分布の密度関数（通常小文字のxでXの特定の実現値を表す）

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{1}$$



正規分布

確率変数の標準化

期待値 $E(X) = \mu$、分散 $V(X) = \sigma^2$ の確率変数 X があるとする。X から期待値を引いて、標準偏差で割って、できた新しい確率変数

$$Z = \frac{X - \mu}{\sigma}$$

は標準化された確率変数で、その期待値 $E(Z) = 0$、分散 $V(Z) = 1$ となる。

分布表から確率を読み取る

- 期待値$\mu = 1$分散$\sigma^2 = 4$の正規分布確率変数Xに関して$X \leq 4.28$の確率を標準正規分布表（テキストp293、付表2）から読み取る。
- まずXを標準化してZとする。Zが標準正規分布（期待値$\mu = 0$分散$\sigma^2 = 1$）に従う確率変数となる。

$$Z = \frac{X - \mu}{\sqrt{\sigma^2}} = \frac{X - 1}{2}$$

- Zに関して標準正規分布表を適用する。$x = 4.28$のとき

$$z = \frac{x - 1}{2} = \frac{4.28 - 1}{2} = 1.64$$

ゆえに、 $P(X \leq 4.28) = P(Z \leq 1.64) \approx 0.95$。

期待値の検定

以下の検定の話が成り立つための前提条件 標本数がかなり大きいまたは母集団が正規分布に従うとする。

仮説 帰無仮説H_0：棄却（否定）したい仮説。
対立仮説H_1：採択し（認め）たい仮説。

例 国民の平均年収 が$20,000$ドルを超えたかどうかを調べる。

H_0：国民の平均年収 が$20,000$ドル以下である；
H_1：国民の平均年収 が$20,000$ドルを超えた。

検定の根拠となる定理

Theorem (中心極限定理)

X_i が平均（期待値）μ と分散 σ^2 を持つ独立同一な分布に従うとき，n が大きくなるにつれ，$\sqrt{n}$ 倍した標準化した $\bar{X}$ が標準正規分布に収束する（近づく）。

$$z = \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} \xrightarrow{d} N(0, 1).$$

検定の根拠となる定理（続き）

Theorem

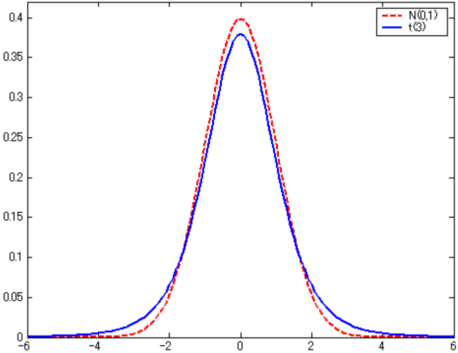
上の定理の条件の下で、σ の代わりに σ の推定量 s（標準偏差、標本分散の平方根）を使った場合、t 値

$$t = \frac{\sqrt{n}(\bar{X} - \mu)}{S} \xrightarrow{d} t_{(n-1)}.$$

となる。ただし、$t_{(n-1)}$ は自由度 $n-1$ の t 分布を表す。

t分布

正規分布は平均と分散によって密度関数が決まるが、t 分布は自由度というパラメーターにより密度関数が決まる。自由度が大きくになるにつれ t 分布が正規分布に近づき、密度関数のグラフが同じようになっていく。下のグラフは標準正規分布の密度関数のグラフ（赤い点線）と自由度が 3 の t 分布の密度関数のグラフ（ブルーの実線）



平均の片側検定

平均の検定には片側検定と両側検定がある、片側検定は平均がある値より大きいかどうか、または小さいかどうかに関して別々に検定する。両側検定の場合、両方を同時に検定する。

平均の片側検定の例題

国民の平均年収 が $20,000$ ドルを超えたかどうかが先進国であるかどうかを判断するための一つおおよそな指標となる。A 国が先進国であるかどうかをこの指標で検討したい。A 国の国民の個人年収を確率変数 X とする。X の期待値（全国民の平均年収）に関して検定を行う。10000 人をくじ引きで選んで（無作為標本)、年収のデータを収集する、その平均を計算して $\bar{X} = 20,200$ ドル、標本標準偏差を計算して $S_x = 7500$ になったとする。期待値 μ が $20,000$ ドルより大きいかどうかを検定する。

基本的な考え方（反証法的な考え方）

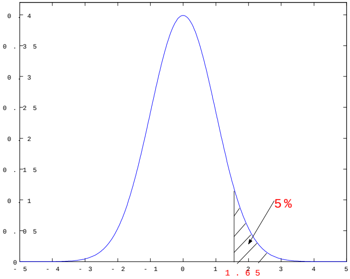
- まずXは平均（期待値）が$\mu \leq 20000$、標準偏差が$\sigma = S_x = 7500$の正規分布$N\left(20000, 7500^2\right)$に従うと仮定する（すなわち帰無仮説は$\mathrm{H}_0$：$\mu \leq 20000$）。
- 以上の仮定の下で、$\bar{X} \geq 20200$の確率を調べる。
- もし以上の仮定の下で$\bar{X} \geq 20200$確率は極めて小さいなら、仮定がデータから得られた結果$\bar{X} = 20200$と矛盾する（仮定が正しいなら、$\bar{X}$は20200という大きい値にならないはず）。ゆえに$\mu \leq 20000$の仮説$\mu \leq 20000$が偽であると判断する（棄却する）。
- もし以上の仮定の下で$\bar{X} \geq 20200$確率は小さくない、仮定がデータから得られた結果$\bar{X} = 20200$と矛盾しない、$\mu \leq 20000$の仮説$\mu \leq 20000$が真であると判断する（採択する）。

有意水準と臨界値（有意水準点）

有意水準 確率が小さいかどうかを判断する基準である。通常αで表す、慣例として1%か5%とする。

臨界値（有意水準点） 確率が有意水準aに対応している確率変数の値。たとえば、今の例では$P(X > c) = a$となるようなcの値。

- 有意水準を$a = 5\%$にした場合、影の部分の面積は$a = 5\%$を表している。その影の左端の座標は臨界値となる。今のグラフの例では臨界値は1.65となる。数式で表すと、$P(X > 1.65) = 5\%$。



検定の手順

1. 有意水準を$\alpha = 5\%$とする。
2. 帰無仮説H_0：期待値$\mu \leq 20000$；対立仮説H_1：期待値$\mu > 20000$とする。
3. 検定統計量t値を計算する。
$t = \sqrt{n}\left(\bar{X} - \mu\right)/s = 100 \times \left(20200 - 20000\right)/7500 = 2.66$
4. t分布表より（自由度が大きい場合標準正規分布表を利用しても良い)、自由度（平均の検定のとき、自由度は$n-1$）が9999のt分布の5%の臨界値（有意水準点）が約1.65である。
5. 検定統計量の値$t = 2.66 > 1.65$であるため[2]帰無仮説H_0：期待値$\mu \leq 20000$が棄却される。国民の平均年収が$20,000$ドルを超えて先進国といえると判断する。

[2] $\bar{X}$がその実現値20100を超える確率が5%よりも小さいと意味する

注

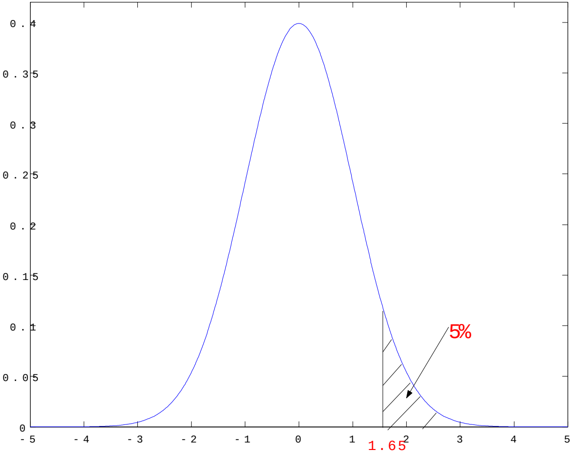
1. 今の例題の対立仮説は$\mu > 20000$、より大きい$(>)$となっている。より小さい（$<$）たとえば対立仮説が$\mu < 20000$となる場合、t値を計算して左側の有意水準点と比較する。t値が左側の有意水準点より小さければ帰無仮説を棄却する、逆の場合だったら帰無仮説を採択する。
2. 左側の有意水準点は分布表から見つかった右側の有意水準点の値$\times(-1)$)、たとえば、右側の5%有意水準点が1.65だったら、左側は-1.65となる。

密度関数のグラフで見る

グラフで示すなら、検定統計量 t が t 分布の裾の端に入って有意水準点よりも右にあって、$\mu \leq 20000$ の仮説が偽であると判断し棄却する。

棄却域 影の部分は棄却域と呼ぶ。検定統計量 t 値が棄却域に入った場合、帰無仮説は棄却される。

1. 対立仮説はより大きい（$>$）となっている場合、棄却域はグラフの右裾にある。t 値が棄却域に入った場合（すなわちは検定統計量 t が有意水準点よりも右にある場合)、帰無仮説を棄却する。
2. 対立仮説はより小さい（$<$）となっている場合、棄却域はグラフの左裾にある。t 値が棄却域に入った場合（すなわちは検定統計量 t が有意水準点よりも左にある場合)、帰無仮説を棄却する。
3. 以下の図は対立仮説はより大きい（$>$）となっている場合の棄却域を示している。


5%
1.65

演習問題

ある美容室が割引サービスを行った、この割引サービスによって、一日の平均来客数が増えたかどうかを調べたい。この美容室の普段の平均来客数が10人。割引サービスを実施後、25日間来客数を集計して平均と標準偏差を計算して$\bar{X} = 12$、$s = 3$だとする。検定を行って一日平均の来客数が増えたかどうかを判断してください。ヒント：H_0：$\mu \leq 10$；H_1：$\mu > 10$。

検定の過誤

	H_0を採択	H_0を棄却H_1を採択
H_0が真	判断が正しい （確率$1-\alpha$で真を真に）	第１種の過誤 （確率αで真を偽に）
H_0が偽	第２種の過誤 （確率βで偽を真と判断）	判断が正しい （確率$1-\beta$で偽を偽と判断）

有意水準 第１種の過誤を犯す確率α。通常検定者が自分で定まる。教科書的には$1\%, 5\%, 10\%$などの値をとる。

検出力 偽を偽と正しく判断する確率$1-\beta$。高いほど良い。検定統計量、たとえばt値の真の分布に依存する。

検定方法の優劣 通常は有意水準αが低く検出力が高い（βが低い）ほうが良いとする。すなわち、２種類の過誤を低く抑えられる検定が良い。

平均の両側検定

部品のサイズの平均の検定を例に説明する

工場から送ってきた大量な同じ種類の部品を検査することを考える。納品の中から100個を無作為に抽出して、直径を図り平均と標準偏差を計算し$\bar{X} = 3.2\,\mathrm{cm}$, $s = 2$となった。良品の条件として直径の期待値が$\mu = 3$と決まっている。$\bar{X}$の値を利用して、$\mu = 3$であるかどうかの検定を考える。

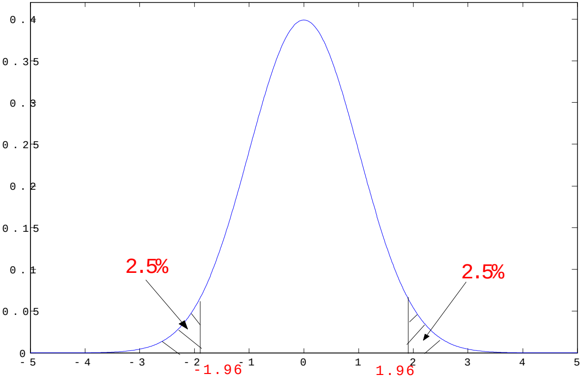
方針 方法は片側検定とほぼ同じである。異なるところは

1. 帰無仮説は$=$を使う、対立仮説は$<, >$両方を使う。たとえば、H_0：$\mu = 3$；H_1：$\mu > 3$または$\mu < 3$。
2. 有意水準を決めてから、それを半分にして、有意水準点の値を調べる。たとえば、有意水準を5%にした場合、それの半分2.5%の有意水準点の値を調べる。
3. 検定統計量t値の計算などは片側のときと同じであるが、最後にtの絶対値と有意水準点の値（たとえば、2.5%の有意水準点の値）と比較する。$|t| >$有意水準点の値であれば、帰無仮説を棄却する、逆に$|t| \leq$有意水準点の値であれば、帰無仮説を採択する。

検定の手順

1. **両側検定を行う。有意水準を**5%**とする。**
2. H_0：$\mu = 3$；H_1：$\mu > 3$**または**$\mu < 3$。
3. $t = \sqrt{n}\left(\bar{X} - \mu\right)/s = 10 \times 0.2/2 = 1$。
4. **分布表より、自由度**99**の**2.5%**有意水準点が**1.96**となる。**
5. $|t| < 1.96$、**帰無仮説は採択される。納品は良品であると判断する。**

- **両側検定の棄却域は両裾の影の部分に対応する。**t**値がどちらに入っても帰無仮説**H_0**は棄却される。**


0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
-5
-4
-3
-2
-1
0
1
2
3
4
5
2.5%
2.5%
-1.96
1.96