**2016 年度秋学期　統計学　第１５回**

# 演習（2）

1. 下の記述は，統計学の観点からみて正しいかどうかを答えよ。正しいときは理由を説明せよ。正しくないときは，どういう点がどのように正しくないかを説明せよ。

   母集団の平均の区間推定を行なう問題で，取り出した標本を使って計算した結果，「母平均の 95 パーセント信頼区間は 20 から 30 である」という結論を得た。このとき，母平均が 20 以上 30 以下である確率は 95 パーセントである。

2. X薬品の「Y」という薬は，１つ１グラムの錠剤となっている。いま，10 個の錠剤を無作為抽出し，各々の錠剤に含まれる物質Ｐの量を調べた。その結果，各錠剤の物質Ｐの量（ミリグラム）は次の通りであった。

   1.2　0.8　0.9　0.9　1.0　1.3　1.2　1.0　0.8　0.9

   (a) 講義で説明した知識を使って，「薬Ｙに含まれる物質Ｐの割合」を区間推定するには，この測定がどのようなものであると仮定できる必要があるか。
   (b) (a) で答えた仮定が正しいとして，信頼係数 95% で (a) の区間推定を行なえ。

3. 10 人の被験者に，薬Ａを与えた場合と薬Ｂを与えた場合とで，それぞれ同じ検査を行うと，次の表の通りとなった。このとき，薬Ｂは薬Ａよりも検査の数値を高くする働きがあるといえるかを，有意水準 5%の $t$ 検定を用いて答えよ。

表 1: 検査結果

| 被験者番号 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| 薬Ａ | 60 | 65 | 50 | 70 | 80 | 40 | 30 | 80 | 50 | 60 |
| 薬Ｂ | 64 | 63 | 48 | 75 | 83 | 38 | 32 | 83 | 53 | 66 |

## 解答例

1. 母平均は，調査した人が知らないだけで，標本抽出には関係なく既に決まっている。だから，「20 以上 30 以下」という具体的区間を求めた段階で，母平均が「20 以上 30 以下」であるかどうかはすでに決まっている。「95% 信頼区間」とは，この方法で信頼区間を何度も求めると，そのうち 95% は母平均を本当に含んでいる区間である，という意味であり，「20 以上 30 以下」はそれらの信頼区間のうちのひとつにすぎない。

2. (a) 薬Yの各錠に含まれる物質Pの割合が，正規分布にしたがっていること。

   (b) 薬Y全体での 1 グラムあたりの物質P の量を $\mu$，抽出された各錠剤での物質P の量の平均を $\bar{X}$，その不偏分散を $s^2$，抽出された錠剤の数（標本サイズ）を $n$ とする。このとき，問題文と (a) の仮定から，各錠剤が含む物質P の量は，平均 $\mu$ の正規分布にしたがうと考えられる。したがって，

$$t = \frac{\bar{X} - \mu}{\sqrt{s^2/n}} \tag{A1}$$

   という値（$t$ 統計量）は自由度 $n-1$ の $t$ 分布 $t(n-1)$ にしたがうので，$t_{0.025}(n-1)$ を $t(n-1)$ の上側 2.5 パーセント点とすると

$$P\left(\bar{X} - t_{0.025}(n-1)\sqrt{\frac{s^2}{n}} \leq \mu \leq \bar{X} + t_{0.025}(n-1)\sqrt{\frac{s^2}{n}}\right) = 0.95 \tag{A2}$$

   となるので，$\mu$ の 95%信頼区間の上限・下限は，(A2) 式のかっこ内の不等式の上限・下限となる。

   $\bar{X} = (1.2+0.8+\cdots+0.8+0.9)/10 = 1.0$，$s^2 = \{(1.2-1.0)^2+\cdots+(0.9-1.0)^2\}/(10-1) = 0.031$，$n = 10$，$t_{0.025}(9) = 2.262$ であるから，$\mu$ の 95%信頼区間は $[0.87, 1.13]$（ミリグラム）となる。したがって，物質P の割合は $[0.087, 0.113](\%)$ となる。

3. 各被験者について，薬B での数値が薬A での数値に比べていくら大きくなっているかを，次の表で表す。

表 A1: 検査結果と「差」

| 被験者番号 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| 薬A | 60 | 65 | 50 | 70 | 80 | 40 | 30 | 80 | 50 | 60 |
| 薬B | 64 | 63 | 48 | 75 | 83 | 38 | 32 | 83 | 53 | 66 |
| 差 | 4 | $-2$ | $-2$ | 5 | 3 | $-2$ | 2 | 3 | 3 | 6 |

この「差」が正規母集団からの標本であると考えて，母集団全体で「差」の平均を $\mu$ とするとき，$\mu$ が 0 より大きいといえるかどうかを検定する。帰無仮説を $H_0 : \mu = 0$，対立仮説を $H_1 : \mu > 0$ とする片側検定を行う。

標本での「差」の平均を $\bar{X}$ ，不偏分散を $s^2$，標本サイズを $n$ とすると，

$$t = \frac{\bar{X} - \mu}{\sqrt{s^2/n}} \tag{A3}$$

という値は，自由度 $n-1$ の $t$ 分布 $t(n-1)$ にしたがう。いま，帰無仮説 $H_0 : \mu = 0$ が正しいとすると，

$$t = \frac{\bar{X}}{\sqrt{s^2/n}} \tag{A4}$$

が自由度 $n-1$ の $t$ 分布 $t(n-1)$ にしたがう。

表の数値から計算すると，「差」の標本平均 $\bar{X}=2.0$，不偏分散 $s^2=8.89$，標本サイズ $n=10$ なので，式 (A4) の $t$ 統計量の値は 2.121 となる。一方，自由度は $n-1=9$ で，このときの上側 5%点は数表より $t_{0.05}(9)=1.8331$ である。$t$ 統計量が上側 5%点より大きいので，帰無仮説は棄却され，母集団全体で「差」の平均が 0 より大きい，すなわち，有意水準 5%で，薬 B は薬 A よりも検査の数値を高くする働きがあるといえる。

（この問題のような検定を「対応のある $t$ 検定」といい，実際の研究でもよく用いられます。）