※p.32～34の解答

1．一般に、「検定に合格」とはどういう状態を指すか。以下の（1）～（8）のそれぞれについて、正しい方に〇を付けなさい。

（1）　A　調査でわかったことが母集団についても成り立つ
　　　B　調査でわかったことが母集団については成り立たない

（2）　A　調査結果が偶然の産物にすぎない可能性が結構ある
　　　B　調査結果が偶然の産物である可能性はほとんどない

（3）　A　帰無仮説を棄却する
　　　B　対立仮説を棄却する

（4）　A　帰無仮説を採択する
　　　B　対立仮説を採択する

（5）　A　手元のような調査結果が偶然に得られる確率が5%以上ある
　　　B　手元のような調査結果が偶然に得られる確率は5%未満しかない

（6）　A　検定統計量が臨界値を超えている（臨界値の外側）
　　　B　検定統計量が臨界値を超えていない（臨界値の内側）

（7）　A　調査結果が統計的に有意である
　　　B　調査結果が統計的に有意ではない

（8）　A　平均の差の検定の場合、「平均の差がある」といえる
　　　B　　　　〃　　　　　　、「平均の差がある」といえない

2．次のような場合に、検定の結果はどのように判断されるのか。（1）～（8）のそれぞれについて、正しい方に〇を付け、結論を具体的な文で記しなさい。

（1）顧客満足度調査では、男性の平均が5.5点、女性の平均が7.2点だった。平均の差を検定すると、t＝-3.27で、臨界値が±2.02。
⇒　検定統計量が臨界値を［超えている・超えていない］
　　検定に［合格・不合格］
　　H_0を［棄却・採択］
　　結論（…といえる/いえないという形で）：

　　母集団で　男女の平均の差があるといえる

（2）去年の4年生は選挙に行ったことのある学生が45%しかいなかったが、今年の4年生は52%いた。今年の方が選挙に行った比率が高いといってよいか、0/1のダミー変数と考えて、比率の差を検定すると、t＝1.44で、臨界値が±1.96。
⇒　検定統計量が臨界値を［超えている・超えていない］
　　検定に［合格・不合格］
　　H_0を［棄却・採択］
　　結論（…といえる/いえないという形で）：

　　母集団で　今年の方が選挙に行った比率が高いとはいえない

（3）小学生のテレビ視聴時間は1日1時間程度と予想していたが、調査では平均1.3時間という結果が出た。平均は1時間でないと結論づけてよいか、平均の検定をすると、t＝1.69で、臨界値が±1.96。

⇒　検定統計量が臨界値を［超えている・超えていない］
検定に［合格・不合格］
H_0を［棄却・採択］
結論（…といえる/いえないという形で）：

母集団で　テレビ視聴時間の平均は1時間でないとはいえない

（4）－5点～＋5点の範囲で、自分の性格を自己評価してもらったところ、平均値は－0.8点になった。性格の自己評価は標準（±0）よりも低いといってよいか。平均の検定の結果は、t＝-4.29で、臨界値が±2.04。

⇒　検定統計量が臨界値を［超えている・超えていない］
検定に［合格・不合格］
H_0を［棄却・採択］
結論（…といえる/いえないという形で）：

母集団で　性格の自己評価が標準（±0）よりも低いといえる

3.「幼児期に母親が外で働くことは子どもの成長に悪い」という意見に賛成か反対かを＋2点（とても賛成）～－2点（とても反対）の5段階で答えてもらう調査を行った。996人から回答を得て、平均値が-0.52、標準偏差が1.25であった。

（1）95％の信頼度で平均得点の信頼区間を区間推定せよ。

自由度996-1=995は十分大きいので、自由度∞のt分布（標準正規分布）を利用する。

$$\bar{X} - 1.96 \times \frac{s}{\sqrt{n}} < \mu < \bar{X} + 1.96 \times \frac{s}{\sqrt{n}}$$

$$-0.52 - 1.96 \times \frac{1.25}{\sqrt{996}} < \mu < -0.52 + 1.96 \times \frac{1.25}{\sqrt{996}}$$

$$-0.52 - 0.08 < \mu < -0.52 + 0.08$$

$$-0.60 < \mu < -0.44$$

母集団の平均得点は-0.60～-0.44と推定できる

（2）母集団でも、「平均値が0より小さい（全体的には、母親の仕事は有害と考えられない傾向がある）」と判断してよいであろうか。有意水準5％で検定せよ。（両側検定で）

H_0：賛否の平均値は0である（$\mu=0$）
H_1：賛否の平均値は0より小さい（$\mu \neq 0$　あるいは　$\mu < 0$）
平均の検定を行う

$$t = \frac{\bar{X}-\mu}{\frac{s}{\sqrt{n}}} = \frac{-0.52-0}{\frac{1.25}{\sqrt{996}}} = -13.13$$

自由度996-1=995のt分布（標準正規分布）で、臨界値は±1.96
臨界値を超えているので、H_1を採択
賛否の平均値は0より小さいといえる

4.　ある2つの地域で5年生を20人ずつ無作為に選んで、算数の学力試験を行った。その結果、地域Aでは、平均得点が73点、標準偏差が10.0であった。一方、地域Bでは、平均得点が68点、標準偏差が8.0であった。
2つの地域で算数の学力に差があると言ってよいであろうか（平均得点に差があるといってよいであろうか）。有意水準5％で検定を行え。（両側検定で）

H_0：2つの地域で平均得点に差がない（$\mu_1 = \mu_2$）
H_1：2つの地域で平均得点に差がある（$\mu_1 \neq \mu_2$）
平均の差の検定を行う

$$t = \frac{\bar{X}_1-\bar{X}_2}{s_{\bar{x}_1-\bar{x}_2}} = \frac{73-68}{2.86} = 1.75 \qquad \left(※ s_{\bar{x}_1-\bar{x}_2} = \sqrt{\frac{19\times10.0^2+19\times8.0^2}{38} \times \left(\frac{1}{20}+\frac{1}{20}\right)} = \sqrt{\frac{10.0^2+8.0^2}{20}} = 2.86\right)$$

自由度20+20-2=38のt分布（自由度40で代用）で、臨界値は±2.02
臨界値を超えていないので、H_0を採択
2つの地域で平均得点に差があるとはいえない

5. 30代の既婚男性と未婚男性を200人ずつ調べると、自分の父親がすでに亡くなっている人の割合に違いがあった。父親の死亡率は、既婚男性では10%だったのに対して、未婚男性では19%であった。

（1） この結果から未婚男性の方が父親の死亡率が高いと一般化してよいだろうか。平均の差の検定を応用して、比率の差を有意水準5%で検定しなさい。

ダミー変数とみなして、それぞれのグループの平均、標準偏差を算出すると、

未婚：$n_1 = 200, \quad \bar{X}_1 = p = 0.10, \quad s_1 = \sqrt{p(1-p)} = \sqrt{0.10 \times 0.90} = 0.30$

既婚：$n_2 = 200, \quad \bar{X}_2 = p = 0.19, \quad s_2 = \sqrt{p(1-p)} = \sqrt{0.19 \times 0.81} = 0.39$

H₀：既婚男性と未婚男性で、父親の死亡率は変わらない（$\mu_1 = \mu_2$）

H₁：既婚男性より未婚男性の方が父親の死亡率が高い（$\mu_1 \neq \mu_2$ あるいは $\mu_1 < \mu_2$）

平均の差の検定の応用で比率の差の検定を行う

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s_{\bar{x}_1 - \bar{x}_2}} = \frac{0.10 - 0.19}{0.035} = -2.57 \left(※ s_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{199 \times 0.30^2 + 199 \times 0.39^2}{398} \times \left(\frac{1}{200} + \frac{1}{200}\right)} = \sqrt{\frac{0.30^2 + 0.39^2}{200}} = 0.035 \right)$$

自由度200+200−2=398のt分布（標準正規分布）で、臨界値は±1.96

臨界値を超えているので、H₁を採択

未婚男性の方が父親の死亡率が高いと一般化してよい

（2） 既婚男性と未婚男性の父親死亡率を、それぞれ95%の信頼度で区間推定しなさい。

自由度200−1=199は十分大きいので、自由度∞のt分布（標準正規分布）を利用する。

$$\bar{X} - 1.96 \times \frac{s}{\sqrt{n}} < \mu < \bar{X} + 1.96 \times \frac{s}{\sqrt{n}}$$

既婚男性の父親死亡率は、以下のとおり6〜14%と推定できる。

$$0.10 - 1.96 \times \frac{0.30}{\sqrt{200}} < \mu < 0.10 - 1.96 \times \frac{0.30}{\sqrt{200}}$$

$$0.10 - 0.04 < \mu < 0.10 + 0.04$$

$$0.06 < \mu < 0.14$$

未婚男性の場合は、同様に14〜24%と推定できる。

$$0.19 - 1.96 \times \frac{0.39}{\sqrt{200}} < \mu < 0.19 - 1.96 \times \frac{0.39}{\sqrt{200}}$$

$$0.19 - 0.05 < \mu < 0.19 + 0.05$$

$$0.14 < \mu < 0.24$$

※p.31 の解答

1. ある会社で社員の中から無作為に選ばれた正規雇用の20人、非正規の12人に、会社への満足度を尋ねて得点化した。その結果、正規雇用の満足度は平均50点、標準偏差10であり、非正規の満足度は平均43点、標準偏差8であった。母集団において正規と非正規の間に満足度に差があるといえるか。有意水準5%で検定せよ。

H₀：正規と非正規で平均満足度に差がない（$\mu_1 = \mu_2$）

H₁：正規と非正規で平均満足度に差がある（$\mu_1 \neq \mu_2$）

平均の差の検定を行う

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s_{\bar{x}_1 - \bar{x}_2}} = \frac{50 - 43}{3.40} = 2.06 \qquad \left(※ s_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{19 \times 10.0^2 + 11 \times 8.0^2}{30} \times \left(\frac{1}{20} + \frac{1}{12}\right)} = 3.40 \right)$$

自由度20+12−2=30のt分布で、臨界値は±2.04

臨界値を超えているので、H₁を採択

正規と非正規で平均満足度に差があるといえる

2．2009年全国家族調査の結果から、30代前半（30～34歳）で働く男女の通勤時間を整理すると次のようになった。男性の方が8.1分通勤時間が長いことがわかる。

男性：228名、平均34.0分、標準偏差23.3分

女性：173名、平均25.9分、標準偏差20.7分

（1）この調査結果ら、母集団でも「男性の方が平均通勤時間が長い」と結論づけてもよいだろうか。有意水準5%で平均の差を検定しなさい。

H_0：男女で平均通学時間に違いはない（$\mu_1 = \mu_2$）
H_1：男性の方が平均通学時間が長い（$\mu_1 \neq \mu_2$　あるいは　$\mu_1 > \mu_2$）

平均の差の検定を行う

$t = \frac{\bar{X}_1 - \bar{X}_2}{s_{\bar{x}_1 - \bar{x}_2}} = \frac{34.0-25.9}{2.24} = 3.62$　$\left(※ s_{\bar{x}_1-\bar{x}_2} = \sqrt{\frac{227\times 23.3^2 + 172\times 20.7^2}{399} \times \left(\frac{1}{228} + \frac{1}{173}\right)} = 2.24\right)$

自由度228+173-2=399のt分布（標準正規分布）で、臨界値は±1.96

臨界値を超えているので、H_1を採択

男性の方が平均通学時間が長いといえる

（2）現代日本社会において、30代前半の男女で通勤時間に差があるのはなぜか。あなたの意見を論理的な文章で示しなさい。

（略）※通常考えるべきことは、家事・保育所の送り迎え等に関わるジェンダー差。

3．大阪府の成人から300人を無作為に選び、現在の内閣を支持するかどうかを尋ねた。結果は55%が支持しているというものであった（300人全員が回答したものとする）。

（1）母集団において内閣の支持率は50%より高いといえるか。有意水準5%で検定せよ。

ダミー変数とみなして、平均、標準偏差を算出すると、

$n = 300,\quad \bar{X} = p = 0.55,\quad s = \sqrt{p(1-p)} = \sqrt{0.55 \times 0.45} = 0.497$

H_0：内閣の支持率は50%である（$\mu = 0.50$）
H_1：内閣の支持率は50%より大きい（$\mu \neq 0.50$　あるいは　$\mu > 0.50$）

平均の検定の応用で比率の検定を行う

$t = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}} = \frac{0.55-0.50}{\frac{0.497}{\sqrt{300}}} = 1.74$

自由度300-1=299のt分布（標準正規分布）で、臨界値は±1.96

臨界値を超えていないので、H_0を採択

内閣の支持率が50%より大きいとはいえない

（2）先月、やはり300人に調査をしたときには、支持率が58%であったという。母集団において、先月と今月の間に差があったといえるか。有意水準5%で検定せよ。

ダミー変数とみなして、先月と今月の平均、標準偏差を算出すると、

先月：$n_1 = 300,\quad \bar{X}_1 = p = 0.58,\quad s_1 = \sqrt{p(1-p)} = \sqrt{0.58 \times 0.42} = 0.494$

今月：$n_2 = 300,\quad \bar{X}_2 = p = 0.55,\quad s_2 = \sqrt{p(1-p)} = \sqrt{0.55 \times 0.45} = 0.497$

H_0：先月と今月の支持率に差はない（$\mu_1 = \mu_2$）
H_1：先月と今月の支持率に差がある（$\mu_1 \neq \mu_2$）

平均の差の検定の応用で比率の差の検定を行う

$t = \frac{\bar{X}_1 - \bar{X}_2}{s_{\bar{x}_1 - \bar{x}_2}} = \frac{0.58-0.55}{0.040} = 0.75$

$$\left(※ s_{\bar{x}_1-\bar{x}_2} = \sqrt{\frac{299 \times 0.494^2 + 299 \times 0.497^2}{598} \times \left(\frac{1}{300} + \frac{1}{300}\right)} = \sqrt{\frac{0.494^2 + 0.497^2}{300}} = 0.040\right)$$

自由度300+300-2=598のt分布（標準正規分布）で、臨界値は±1.96

臨界値を超えていないので、H_0を採択

先月と今月の支持率に差があったとはいえない