経済統計：前期期末試験

村澤 康友

2009 年 8 月 5 日

注意： 3 問とも解答すること．

1. （20 点）以下の用語の定義を式または言葉で書きなさい（各 20 字程度）．
   (a) 単純仮説 (b) 検定 (c) 第 1 種の誤り (d) 検出力
2. （30 点）ゴルトンは身長の遺伝を研究した．父親と息子の身長の 2 変量データを (X_i, Y_i)，その身長差を $Z_i := X_i - Y_i$ とする（単位はインチ）．$Z_i \sim \mathrm{N}\left(\mu, \sigma^2\right)$ と仮定する．次の両側検定問題を考える．

$$H_0 : \mu = 0 \quad \text{vs.} \quad H_1 : \mu \neq 0.$$

   ゴルトンのデータの分析結果は以下の通りであった（無作為標本を仮定）．

```
Null hypothesis: population mean = 0
Sample size: n = 928
Sample mean = 0.21972, std. deviation = 2.32494
Test statistic: t(927) = (0.21972 - 0)/0.07632 = 2.87893
Two-tailed p-value = 0.004082
(one-tailed = 0.002041)
```

   (a) 父親と息子の身長差の標本平均・標本分散の値を書きなさい．
   (b) 検定統計量の値を書きなさい．それは H_0 の下でどのような分布にしたがうか？
   (c) この検定問題に対する p 値はどちらか？また有意水準 5 %の検定の結果を判定しなさい．
3. （50 点）府大生の（1 日平均）勉強時間の分布を男女で比較したい．以下の通り記号を定義する．

	男子	女子
母集団分布	$\mathrm{N}\left(\mu_X, \sigma_X^2\right)$	$\mathrm{N}\left(\mu_Y, \sigma_Y^2\right)$
標本	$(X_1, \ldots, X_m)$	$(Y_1, \ldots, Y_n)$
標本平均	$\bar{X}$	$\bar{Y}$
標本分散	s_X^2	s_Y^2

   すべての母数は未知とし（$\sigma_X^2 = \sigma_Y^2$ も仮定できない），無作為標本を仮定する．
   (a)「女子の方が勉強する」という仮説を確かめる検定問題を定式化しなさい．
   (b) $\bar{X} - \bar{Y}$ の分布を求めなさい．
   (c) この検定問題に対する検定統計量を定義しなさい．
   (d) 有意水準 5 %の検定の棄却域を定めなさい．
   (e) 検定統計量の値が -2.00 のとき，この検定問題に対する p 値を求めなさい．

解答例

1. 仮説検定の基本用語
   (a) ただ 1 つの分布を許容する仮説.
      - 「ただ 1 点の母数を許容する仮説」でも OK.
   (b) 統計的仮説の真偽を標本から判定すること.
      - 「標本から」がなければ 0 点.
   (c) H_0 が真なのに H_0 を棄却する誤り.
   (d) 第 2 種の誤りを起こさない確率.
      - 「起こりにくさ」は曖昧な表現なので 2 点.
2. gretl の出力の見方（対標本の平均の差の検定）
   (a) $\bar{Z} = .21972$, $s_Z^2 = 2.32494^2 = 5.405\ldots$.
      - 各 5 点.
   (b) $t = 2.87893$, 自由度 927 の t 分布.
      - 各 5 点.
      - 自由度なしは 0 点.
   (c) 両側検定なので $p = .004082$. $p \le .05$ より H_0 は棄却（一部の息子がまだ成長中？）.
      - 各 5 点.
3. 2 標本の平均の差の検定
   (a)
   $$H_0 : \mu_X = \mu_Y \quad \text{vs.} \quad H_1 : \mu_X < \mu_Y.$$
      - H_0 は $\mu_X \ge \mu_Y$ でも OK.
   (b)
   $$\begin{aligned}\bar{X} &\sim \mathrm{N}\left(\mu_X, \frac{\sigma_X^2}{m}\right),\\ \bar{Y} &\sim \mathrm{N}\left(\mu_Y, \frac{\sigma_Y^2}{n}\right).\end{aligned}$$
   したがって
   $$\bar{X} - \bar{Y} \sim \mathrm{N}\left(\mu_X - \mu_Y, \frac{\sigma_X^2}{m} + \frac{\sigma_Y^2}{n}\right).$$
      - $\bar{X}, \bar{Y}$ の分布で 5 点.
   (c)
   $$Z := \frac{\bar{X} - \bar{Y}}{\sqrt{s_X^2/m + s_Y^2/n}}.$$
      - 未知母数を含むのは 0 点.
   (d) H_0 の下で $Z \overset{a}{\sim} \mathrm{N}(0,1)$. したがって棄却域は $(-\infty, -1.65]$.
   (e) $Z \sim \mathrm{N}(0,1)$ なら
   $$\begin{aligned}p &:= \Pr[Z \le -2]\\ &= \Pr[Z \ge 2]\\ &\approx .02275.\end{aligned}$$

- 負の確率は 0 点.

※答案は返却します．採点や成績に関する質問にも応じます．**オフィスアワーの時間に**研究室まで来てください（夏季休業中は随時）.