中級統計学：復習テスト 23

学籍番号＿＿＿＿＿＿＿＿＿＿氏名＿＿＿＿＿＿＿＿＿＿

2024 年 1 月 15 日

注意：すべての質問に解答しなければ提出とは認めない．正答に修正した上で，復習テスト 21〜26 を順に重ねて左上でホチキス止めし，定期試験実施日（1 月 26 日の予定）に提出すること．

1. $\mathrm{Bin}(1,p)$ から抽出した大きさ n の無作為標本の標本比率（＝標本平均）を $\hat{p}$ とする．次の両側検定問題を考える．

$$H_0 : p = p_0 \quad \text{vs} \quad H_1 : p \neq p_0$$

(a) $\mathrm{Bin}(1,p)$ の平均と分散を求めなさい．

(b) $\hat{p}$ の平均と分散を求めなさい．

(c) $\hat{p}$ の漸近分布を求めなさい．

(d) H_0 の下で漸近的に $\mathrm{N}(0,1)$ にしたがう検定統計量を与えなさい．

(e) H_0 の下で漸近的に $\chi^2(1)$ にしたがう検定統計量を与えなさい．

2. 母集団分布の cdf を $F(.)$ とする．適合度検定問題は

$$H_0 : F(.) = F_0(.) \quad \text{vs} \quad H_1 : F(.) \neq F_0(.)$$

k 階級に分割して分布を表すと

階級	$F(.)$	$F_0(.)$
1	p_1	$p_{0,1}$
$\vdots$	$\vdots$	$\vdots$
k	p_k	$p_{0,k}$
計	1	1

大きさ n の無作為標本における第 j 階級の観測度数を N_j とする．

(a) 観測度数と期待度数を用いてピアソンの χ^2 適合度検定統計量を定義しなさい．

(b) χ^2 適合度検定統計量が次のようにも書けることを示しなさい．

$$\chi^2 = \sum_{j=1}^{k} \frac{n(\hat{p}_j - p_{0,j})^2}{p_{0,j}}$$

(c) $k = 2$ のとき χ^2 適合度検定統計量が母比率の検定統計量と一致することを示しなさい（ヒント：$p_1 + p_2 = 1$）．

解答例

1.（a）$X \sim \mathrm{Bin}(1,p)$ とすると

$$\begin{aligned}
\mathrm{E}(X) &:= 1\cdot p + 0\cdot(1-p)\\
&= p\\
\mathrm{var}(X) &= \mathrm{E}\left(X^2\right) - \mathrm{E}(X)^2\\
&= \mathrm{E}(X) - \mathrm{E}(X)^2\\
&= p - p^2\\
&= p(1-p)
\end{aligned}$$

（b）期待値の線形性より

$$\begin{aligned}
\mathrm{E}(\hat{p}) &= \mathrm{E}\left(\frac{X_1+\cdots+X_n}{n}\right)\\
&= \frac{\mathrm{E}(X_1)+\cdots+\mathrm{E}(X_n)}{n}\\
&= \frac{np}{n}\\
&= p
\end{aligned}$$

$X_1,\dots,X_n$ は独立なので

$$\begin{aligned}
\mathrm{var}(\hat{p}) &= \mathrm{var}\left(\frac{X_1+\cdots+X_n}{n}\right)\\
&= \frac{\mathrm{var}(X_1)+\cdots+\mathrm{var}(X_n)}{n^2}\\
&= \frac{np(1-p)}{n^2}\\
&= \frac{p(1-p)}{n}
\end{aligned}$$

（c）中心極限定理より

$$\hat{p} \overset{a}{\sim} \mathrm{N}\left(p, \frac{p(1-p)}{n}\right)$$

（d）標準化すると

$$\frac{\hat{p}-p}{\sqrt{p(1-p)/n}} \overset{a}{\sim} \mathrm{N}(0,1)$$

$H_0 : p = p_0$ を代入すると，検定統計量は

$$Z := \frac{\hat{p}-p_0}{\sqrt{p_0(1-p_0)/n}}$$

（e）$\mathrm{N}(0,1)$ の 2 乗は $\chi^2(1)$ なので

$$Z^2 = \frac{n(\hat{p}-p_0)^2}{p_0(1-p_0)}$$

2.（a）ピアソンの χ^2 適合度検定統計量は

$$\chi^2 := \sum_{j=1}^{k} \frac{(N_j - np_{0,j})^2}{np_{0,j}}$$

（b）$\hat{p}_j := N_j/n$ より

$$\begin{aligned}
\chi^2 &= \sum_{j=1}^{k} \frac{(n\hat{p}_j - np_{0,j})^2}{np_{0,j}} \\
&= \sum_{j=1}^{k} \frac{n(\hat{p}_j - p_{0,j})^2}{p_{0,j}}
\end{aligned}$$

（c）$k = 2$ なら

$$\begin{aligned}
\chi^2 &= \frac{n(\hat{p}_1 - p_{0,1})^2}{p_{0,1}} + \frac{n(\hat{p}_2 - p_{0,2})^2}{p_{0,2}} \\
&= \frac{n(\hat{p}_1 - p_{0,1})^2}{p_{0,1}} + \frac{n[(1 - \hat{p}_1) - (1 - p_{0,1})]^2}{1 - p_{0,1}} \\
&= \frac{n(\hat{p}_1 - p_{0,1})^2}{p_{0,1}} + \frac{n[-(\hat{p}_1 - p_{0,1})]^2}{1 - p_{0,1}} \\
&= \frac{n(\hat{p}_1 - p_{0,1})^2}{p_{0,1}} + \frac{n(\hat{p}_1 - p_{0,1})^2}{1 - p_{0,1}} \\
&= \frac{(1 - p_{0,1})n(\hat{p}_1 - p_{0,1})^2 + p_{0,1}n(\hat{p}_1 - p_{0,1})^2}{p_{0,1}(1 - p_{0,1})} \\
&= \frac{[(1 - p_{0,1}) + p_{0,1}]n(\hat{p}_1 - p_{0,1})^2}{p_{0,1}(1 - p_{0,1})} \\
&= \frac{n(\hat{p}_1 - p_{0,1})^2}{p_{0,1}(1 - p_{0,1})}
\end{aligned}$$