RSS Higher Certificate in Statistics, Specimen B

Module 3: Basic Statistical Models

Solutions

Question 1

(i) もしホームチームのスコアがシーズンを通して一定の平均値でランダムに起きるとみなすことができるとすれば，ゴール数を説明するのにポアソン分布は妥当な仮定であろう．もしそうであるならば，試合におけるホームチームのゴール数はパラメータ（平均）がこの一定の平均値に等しいポアソン分布に従い，このパラメータを μ とする．

(ii) $\bar{r} = 634/380 = 1.6684$

$$s^2 = \frac{1}{\sum f - 1}\left\{\sum fr^2 - \frac{\left(\sum fr\right)^2}{\sum f}\right\} = \frac{1}{379}\left(1778 - \frac{634^2}{380}\right) = 1.9003$$

(iii) μ を 1.6684 とする．このとき $P(R=0) = e^{-1.6684} = 0.1885$，そして $r=0$ に対する期待度数は $380 \times 0.1885 = 71.65$ となる．

同様に $P(R=1) = 1.6684e^{-1.6684} = 0.3145$， $r=1$ に対する期待度数は 119.51．

したがって観測度数と期待度数は以下のようにまとめられる（問題文から残りの期待度数が得られる）．

r	0	1	2	3	4	≥5	Total
Observed	81	112	101	44	28	14	380
Expected	71.65	119.51	99.72	55.46	23.13	10.51	379.98

（注：期待度数を求める際にとても小さな丸め誤差が生じている．）

検定統計量は

$$X^2 = \sum \frac{(O-E)^2}{E} = \frac{(81-71.65)^2}{71.65} + \frac{(112-119.51)^2}{119.51} + \cdots + \frac{(14-10.51)^2}{10.51} = 6.261$$

であり，これを χ_4^2 と比較する（6 つのセルと 1 つの推定されたパラメータから自由度は 4 となる）．これは有意ではない（5%点は 9.49）．したがって帰無仮説は棄却することができない．すなわち，これらのデータがポアソンモデルに従っていることに反するという根拠はない．

検定では期待度数があまり小さすぎてはいけない．(5 以上がしばしば基準として用いられる．) もし大きな r における度数が併合されないならば，これはそのようなケースにはならなかったであろう．

Question 2

(i) 標本分散は

$$s^2 = \frac{1}{59}\left(4054484 - \frac{15568^2}{60}\right) = \frac{15106.93}{59} = 256.05$$

である．検定される帰無仮説は$\sigma^2 = 256$であり，明らかに（たとえサンプルサイズがかなり小さいとしても）帰無仮説が棄却されそうにないように見える．検定を続けると，検定統計量は

$$\frac{(n-1)s^2}{\sigma^2} = \frac{59 \times 256.05}{256} = 59.01$$

となる．これをχ^2_{59}と比較する．上側 5%点は約 78 であるので帰無仮説を棄却することはできず，分散が 256 ではないという根拠は示されない．

(ii) $\bar{x} = \frac{15568}{60} = 259.46$であり，帰無仮説を$\mu = 266$として検定を行いたい．$\sigma$の値を 16 として（ (i) から高い妥当性があるであろう），以下のような検定統計量．

$$\frac{\bar{x} - 266}{\frac{16}{\sqrt{60}}} = -3.16$$

を用い，標準正規分布 N(0, 1) と比較する．これは N (0, 1) の両側 1%点からかなり離れているので，帰無仮説に反するという強い根拠が示される．したがって，この集団は普通の母集団とは異なった平均値を持つと結論付けるのが理にかなっている．つまり妊娠期間が短いようである．また，N(0, 1) を用いると，$\Phi(-3.16) = 0.0008$が得られ，P-値は0.0016となる．

（**別解**として，標本分散 s^2 (=256.05) を用いることができる．その場合は $(\bar{x} - 266)/(s/\sqrt{60})$ と t_{59} を比較する．ここでは実際にはどちらの方法もほとんど差はない.）

Question 3

(i) 2 つの標本が同じ分散で異なる平均（たいていの場合，帰無仮説は 2 つの平均値を等しいとする）となる独立な正規分布から得られるものでなければならない．標本は無作為抽出であり，互いに独立である．

(ii) $n_1 = 7,\ \bar{x}_1 = 54.56,\ s_1^2 = 534.6956,\ n_2 = 14,\ \bar{x}_2 = 49.56,\ s_2^2 = 261.3001$ であり，またプールされた分散の推定値は $s^2 = \dfrac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 - 1 + n_2 - 1} = 347.6355$ である．

μ_1 と μ_2 をそれぞれ 2 つの町の平均価格としたとき，帰無仮説 $\mu_1 - \mu_2 = 0$ を検定する検定統計量は

$$\frac{\bar{x}_1 - \bar{x}_2 - 0}{s\sqrt{\dfrac{1}{7} + \dfrac{1}{14}}} = \frac{5.27}{8.631} = 0.611$$

である．これを t_{19} と比較すると有意ではないので，帰無仮説は棄却することができない．したがって真の平均が 2 つの町で異なるという根拠はない．

(iii) Town 1 は人口がより多く，住宅のすべてのタイプの数がより多いが，標本サイズは Town 2 の半分しかない．したがって 2 つの町で抽出される確率はかなり異なる．また，標本は 2 つの町の不動産屋のローカル誌の住宅広告に制限されている．たとえ典型的な標本を抽出したとしても，無作為性の前提は疑問である．このようなデータに基づく検定は，たとえ正規性と等分散性が受け入れられたとしても，実用的な理由で疑わしいと考えるべきである．（母分散が等しいことを示すための通常の $F_{6,13}$ 検定では検定統計量の値は 2.05 であり，有意ではない．）

Question 4

(i) カイ 2 乗検定によって「Commercial」と「Purchase」との間で関係がないという帰無仮説を関係があるという対立仮説に対して調べる．観測された度数と帰無仮説の下での期待度数（カッコ内の数値）は以下のようである．

		Commercial		
		A	*B*	*Total*
Purchase	*No*	70 (75)	80 (75)	150
	Yes	30 (25)	20 (25)	50
	Total	100	100	200

検定統計量$=\dfrac{(70-75)^2}{75}+\dfrac{(80-75)^2}{75}+\dfrac{(30-25)^2}{25}+\dfrac{(20-75)^2}{25}=2.667$ である．

（もし Yates の補正を施して $(O-E)^2$ を $(4.5)^2$ とすれば検定統計量は 2.16 となる．）

χ_1^2 と比較すると有意ではないので，「Commercial」と「Purchase」の間で関係があるという根拠はない．よって宣伝のコストや顧客数といった非統計的な根拠によって決定がなされていると考えることができる．

(ii) 対データにおけるマクネマー検定は 2 つの分類間に関連があるかどうかに関する仮説を検定する．この場合の分類はメーカーが広告に用いた媒体（テレビ，インターネット）である．マクネマー検定では「No – No」や「Yes – Yes」の組み合わせとなるメーカーの情報はどちらも用いない．検定統計量$=\dfrac{(5-15)^2}{5+15}=\dfrac{100}{20}=5.00$ となるので，χ_1^2 と比較すると 5%有意である．したがって，宣伝するメディアに違いがないという帰無仮説に反する根拠となる。

(iii) (i)では反応（responses）の 2 種類の異なる無作為標本が得られ，反応に対する比率（たとえば Commercial に yes とする比率と Purchase に yes とする比率）の比較が問題となった．さらなる例として，「男性」と「女性」という性別のグループや「若者」と「年配者」という年齢のグループを区別したときの意見の調査を行うというような状況があげられる．これは同質な母集団から 2 つの標本を抽出してカイ 2 乗検定を行う場合にしばしば起こる状況である．また臨床実験でも応用され，異なったグループの患者（例えば，喫煙者と非喫煙者）が特定の病気を持っている人と持っていない人とにクラス分けされる．

ところが，(ii) は 2 つの独立した標本ではない．このような状況は臨床実験でより多く見られるかもしれない．例えば，2 種類の薬が慢性的な病気を治療するのに用いられるとすると，短期間でいずれかを同じ患者に対して（または少なくとも年齢，性別，総合的な医療面での条件が似ているようなペアの患者に対して）実験を行う．薬が両方とも効いたか，あるいは効かなかったか患者（またはペアの患者）からは情報を得られな

い．マクネマー検定はどちらか一つだけが効いたかという患者について，薬 A が薬 B より効果的かどうかを調べる．これは，データの「Yes - No」「No – Yes」の部分の割合が比較される． (ii) の例では同じメーカーでの実験なので，マクネマー検定が妥当である．