2008年度　政策特論（ノンパラメトリック統計学）
第3回宿題解答例

2009年1月23日

Q. 4.2

超過駐車時間 x が平均20の負の指数分布に従うという仮説より、その分布関数は $F(x) = 1 - e^{-x/20},\ 0 \le x < \infty$ となる。一方、経験分布関数を $S(x)$ とすると、標本データの各値における $F(x),\ S(x)$ は下表のとおりである。

x_i	$F(x_i)$	$S(x_i)$	$F(x_i) - S(x_i)$	$F(x_i) - S(x_{i-1})$
5	0.2212	0.0625	0.1587	0.2212
8	0.3297	0.1250	0.2047	0.2672
10	0.3935	0.1875	0.2060	0.2685
11	0.4231	0.3125	0.1106	0.2356
15	0.5276	0.4375	0.0901	0.2151
17	0.5726	0.5000	0.0726	0.1351
20	0.6321	0.5625	0.0696	0.1321
23	0.6834	0.6250	0.0584	0.1209
29	0.7654	0.6875	0.0779	0.1404
32	0.7981	0.7500	0.0481	0.1106
36	0.8347	0.8125	0.0222	0.0847
42	0.8775	0.8750	0.0025	0.0650
63	0.9571	0.9375	0.0196	0.0821
145	0.9993	1.0000	-0.0007	0.0618

これより、

$$\begin{aligned}\max|F(x_i) - S(x_i)| &= 0.2060 \\ \max|F(x_i) - S(x_{i-1})| &= 0.2685 = D_{16}\end{aligned}$$

Kolmogorov=Smirnov 検定は、既知の分布関数 $F(x)$ に対し

$$\mathrm{H}_0 :\ S(x) = F(x) \ \text{ vs }\ \mathrm{H}_1 :\ S(x) \neq F(x) \text{ for some } x$$

という仮説検定をおこなうものである。Kolmogorov=Smirnov 検定の数表から

$$\mathrm{P}(D_{16} \geq 0.2685) > \mathrm{P}(D_{16} \geq 0.327) = 0.05$$

が得られる。よって、有意水準 10%でも帰無仮説を棄却できない。したがって、超過駐車時間 x が平均 20 の負の指数分布に従うという仮説を置いても問題はない。　■

Q. 4.8

サンプルサイズは $n = 17$ で奇数であるので、$m = (17-1)/2 = 8$。よって $m+1 = 9$ 番目のサンプルを除外して、

$$x_{m+2} - x_1,\ x_{m+3} - x_2,\ \ldots,\ x_{17} - x_8$$

を計算すると、

$$8,\ 19,\ 10,\ 25,\ 1,\ 11,\ 20,\ -3$$

となる。

計算された m 個のデータのうち、負値のものの個数を Y とすると、タイムトレンドがないという帰無仮説の下では $Y \sim B(8, 1/2)$ となる。今のケースでは $Y = 1$ であるから、

$$\mathrm{P}(Y \leq 1) =_8 C_1 \left(\frac{1}{2}\right)^8 +_8 C_0 \left(\frac{1}{2}\right)^8 \doteq 0.0351$$

故に有意水準を 5%とすると、両側検定で帰無仮説を棄却できない。よってタイムトレンドがないという仮説は否定されない。　■

Q. 4.12

元のデータは

76	92	105	86	91	81	103	92
71	132	71	57	48	63	43	60

これを $sign(\boldsymbol{x}_i - \boldsymbol{x}_{i-1})$　　$i = 2, \cdots, 16$ に書きかえると

$$+ + - + - + - - + - - - + - +$$

となる。符合を用いた連の検定ができるようになる。

text の定義に従ってまとめると

N = 15,
m = 7　(正値の数),　n = 8　(負位の数),
R = 11　(連の数),　S = 5

であるから、数表より $\Pr(R \geq 11) = 0.133$。よって有意水準 5% でランダムであるという仮説を棄却できない。

note:$\Pr(R \geq 11)$ **の計算**

$$\begin{aligned}
\Pr(R=15) &= \frac{{}_{7-1}C_{7-1} \times {}_{8-1}C_7 + {}_{7-1}C_7 \times {}_{8-1}C_{7-1}}{{}_{15}C_7} = \frac{1}{6435} = 0.00016 \\
\Pr(R=14) &= 2 \times \frac{{}_{7-1}C_{7-1} \times {}_{8-1}C_{7-1}}{{}_{15}C_7} = \frac{2 \times 7}{6435} = 0.00228 \\
\Pr(R=13) &= \frac{{}_{7-1}C_{6-1} \times {}_{8-1}C_6 + {}_{7-1}C_6 \times {}_{8-1}C_{6-1}}{{}_{15}C_7} = \frac{42+21}{6435} = 0.00979 \\
\Pr(R=12) &= 2 \times \frac{{}_{7-1}C_{6-1} \times {}_{8-1}C_{6-1}}{{}_{15}C_7} = \frac{2 \times 126}{6435} = 0.03916 \\
\Pr(R=11) &= \frac{{}_{7-1}C_{5-1} \times {}_{8-1}C_5 + {}_{7-1}C_5 \times {}_{8-1}C_{5-1}}{{}_{15}C_7} = \frac{15 \times 21 + 6 \times 35}{6435} = 0.08159
\end{aligned}$$

以上より、$\Pr(R \geq 11) = 0.13287$。 ■