

政治学方法論 II 課題 1

宋財沄 (SONG Jaehyun) (123J009J)

2015-4-10(金)

1 問題 1

p 値の意味を説明しなさい。

p 値とは帰無仮説が正しいと仮定したとき、現在の推定値あるいは 0 からそれ以上のかげ離れた推定値が得られる確率を意味する。例えば、 $\beta_0 = 0$ を帰無仮説とし、推定から $\beta_0 = 1.0$ で $p = 0.05$ という結果が得られた場合、 $P(\beta_0 \geq 1.0) = 0.05$ ということの意味する。

2 問題 2

「有意確率が 5% 以下なので帰無仮説は正しくない。」という文の問題点を指摘し、正しい文（または文章）に直しなさい。

有意確率を p 値と直すべきである。また、分析の前の段階で棄却域を予め設定し、 p 値がそれを下回る場合に、帰無仮説を棄却する。つまり、帰無仮説の棄却は恣意的な基準で行われるため、帰無仮説が正しくないとは断言することは出来ない。したがって、以上の文は「 p 値が予め設定した棄却域、0.05 を下回るため、帰無仮説は棄却された（あるいは、仮説は支持された）。」と直すべきである。

3 問題 3

人口の 1% が罹る病気に対する検査を考える。陽性の者が検査を受けたとき、97% の確率で「陽性」を示し、陰性の者が検査を受けたときは 95% の確率で「陰性」を示すとする。

(1) ある者の 1 回目の検査結果が陽性のとき、この病気に罹っている確率を求めよ。

$P(D) = 0.01$: 病気に罹っている確率

$P(\neg D) = 0.99$: 病気に罹っていない確率

$P(Pos|D) = 0.97$: 病気に罹っている時に検査から「陽性」が出る確率

$P(\neg Pos|D) = 0.03$: 病気に罹っている時に検査から「陰性」が出る確率

$P(Pos|\neg D) = 0.05$: 病気に罹っていない時に検査から「陽性」が出る確率

$P(\neg Pos|\neg D) = 0.95$: 病気に罹っていない時に検査から「陰性」が出る確率

$$\begin{aligned} P(D|Pos) &= \frac{P(Pos|D)P(D)}{P(Pos)} \\ &= \frac{0.97 \times 0.01 = 0.0097}{P(Pos)} \\ P(Pos) &= P(Pos|D)P(D) + P(Pos|\neg D)P(\neg D) \\ &= 0.97 \times 0.01 + 0.05 \times 0.99 = 0.0592 \\ P(D|Pos) &= \frac{0.0097}{0.0592} \\ &= 0.16385\dots \end{aligned} \tag{1}$$

1 回目の検査結果が陽性のとき、この病気に罹っている確率は約 16.38% である。

(2) 1 回目の検査結果が陽性で、2 回目の検査結果も陽性のとき、この病気に罹っている確率を求めよ。

問 3-(1) の設定の内、 $P(D) = 0.01$ を $P(D) = 0.163851$ に、 $P(\neg D) = 0.99$ を $P(\neg D) = 0.836149$ に更新する。

$$\begin{aligned}
P(D|Pos) &= \frac{P(Pos|D)P(D)}{P(Pos)} \\
&= \frac{0.97 \times 0.163851}{0.200743} = 0.158936 \\
P(Pos) &= P(Pos|D)P(D) + P(Pos|\neg D)P(\neg D) \\
&= 0.97 \times 0.163851 + 0.05 \times 0.836149 = 0.200743 \\
P(D|Pos) &= \frac{0.158936}{0.200743} \\
&= 0.791739... \tag{2}
\end{aligned}$$

1回目の検査結果が陽性で、2回目の検査結果も陽性のとき、この病気に罹っている確率は約79.17%である。

- (3) 3回検査を受けたところ、1回目は陽性、2回目も陽性、3回目は陰性だった。このとき、この病気に罹っている確率を求めよ。

問3-(1)の設定の内、 $P(D) = 0.01$ を $P(D) = 0.791739$ に、 $P(\neg D) = 0.99$ を $P(\neg D) = 0.208261$ に更新する。

$$\begin{aligned}
P(D|\neg Pos) &= \frac{P(\neg Pos|D)P(D)}{P(\neg Pos)} \\
&= \frac{0.03 \times 0.791739}{0.22136} = 0.107185... \\
P(\neg Pos) &= P(\neg Pos|D)P(D) + P(\neg Pos|\neg D)P(\neg D) \\
&= 0.03 \times 0.791739 + 0.95 \times 0.208261 = 0.22136 \\
P(D|\neg Pos) &= \frac{0.02376}{0.22136} \\
&= 0.107185... \tag{3}
\end{aligned}$$

3回検査を受けたところ、1回目は陽性、2回目も陽性、3回目は陰性だった。このとき、この病気に罹っている確率は約10.72%である。

- (4) 3回検査を受けたところ、1回目は陰性、2回目は陽性、3回目は陽性だった。このとき、この病気に罹っている確率を求めよ。

1回目の陰性で病気に罹っている確率は、

$$\begin{aligned}
P(D|\neg Pos) &= \frac{P(\neg Pos|D)P(D)}{P(\neg Pos)} \\
&= \frac{0.03 \times 0.01 = 0.0003}{P(\neg Pos)} \\
P(\neg Pos) &= P(\neg Pos|D)P(D) + P(\neg Pos|\neg D)P(\neg D) \\
&= 0.93 \times 0.01 + 0.95 \times 0.99 = 0.9408 \\
P(D|Pos) &= \frac{0.0003}{0.9408} \\
&= 0.000319... \tag{4}
\end{aligned}$$

であり、次に $P(D) = 0.01$ を $P(D) = 0.000319$ に、 $P(\neg D) = 0.99$ を $P(\neg D) = 0.999681$ に更新する。2回目の陽性で病気に罹っている確率は、

$$\begin{aligned}
P(D|Pos) &= \frac{P(Pos|D)P(D)}{P(Pos)} \\
&= \frac{0.97 \times 0.000319 = 0.00030943}{P(Pos)} \\
P(Pos) &= P(Pos|D)P(D) + P(Pos|\neg D)P(\neg D) \\
&= 0.97 \times 0.000319 + 0.05 \times 0.999681 = 0.05029348 \\
P(D|Pos) &= \frac{0.00030943}{0.05029348} \\
&= 0.006152... \tag{5}
\end{aligned}$$

である。続いて $P(D)$ を 0.006152 に、 $P(\neg D)$ を 0.993848 に更新し、3回目の陽性で病気に罹っている確率を求める。

$$\begin{aligned}
P(D|Pos) &= \frac{P(Pos|D)P(D)}{P(Pos)} \\
&= \frac{0.97 \times 0.006152 = 0.00596744}{P(Pos)} \\
P(Pos) &= P(Pos|D)P(D) + P(Pos|\neg D)P(\neg D) \\
&= 0.97 \times 0.006152 + 0.05 \times 0.993848 = 0.05565984 \\
P(D|Pos) &= \frac{0.00596744}{0.05565984} \\
&= 0.107213... \tag{6}
\end{aligned}$$

最終的にこの病気に罹っている確率は約 10.72% である。

(5) (3) と (4) の結果はどのようなことを示唆するか述べよ。

問 3-3 の結果は約 0.107185、問 3-4 は約 0.107213 であり、若干の違いはあるが、これは小数点の丸めによるものだと考えられる。したがって、エクセルを用いて二つの計算結果を厳密に比較する。

表 1 問 3-3 の結果

	Result	$P(D)$	$P(\neg D)$	$P(Pos)$	$P(\neg Pos)$	$P(D Result)$
1 回目	Pos	0.01	0.99	0.059200	0.940800	0.163851
2 回目	Pos	0.163851	0.836149	0.200743	0.799257	0.791737
3 回目	$\neg Pos$	0.791737	0.208263	0.778398	0.221602	0.107184

表 2 問 3-4 の結果

	Result	$P(D)$	$P(\neg D)$	$P(Pos)$	$P(\neg Pos)$	$P(D Result)$
1 回目	$\neg Pos$	0.01	0.99	0.059200	0.940800	0.000319
2 回目	Pos	0.000319	0.999681	0.050293	0.949707	0.006150
3 回目	Pos	0.006150	0.993850	0.055658	0.944342	0.107184

表 1、2 は問 3-3 と 3-4 の計算過程であり、 $P(Pos|D)$ 、 $P(\neg Pos|\neg D)$ は変化しないため省略した。各表の右下のセルを見ると結果が一致することが確認できる。つまり、検査の結果の順番とは関係なく同じ結果が得られると言える。