

付録

SPSS によるベイズ統計の手順

●—— 2つのグループの差の検定の場合

ベイズ統計の出発点は、次のベイズの定理です。

ベイズの定理

$$Pr(A|B) = \frac{Pr(B|A)}{Pr(B)} \times Pr(A)$$

ただし、

- $Pr(A)$ …… 事象 A の起こる確率
- $Pr(A|B)$ …… 事象 B が起こったという条件のもとで
事象 A が起こる確率
- $Pr(B)$ …… 事象 B の起こる確率

条件付確率のこと



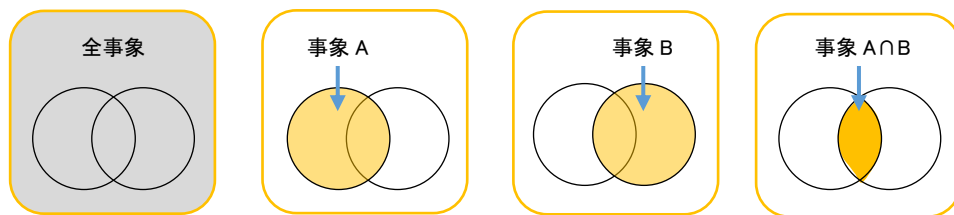
このとき、

- $Pr(A)$ …… 事前確率
- $Pr(A|B)$ …… 条件付確率

といいます。

このベイズの定理は、一見、神秘的に見えるのですが、……

2つの事象 A , B が, 次のようになっているとします.



このとき, 次の等号が成り立つのは, ごく自然です.

$$\frac{Pr(A \cap B)}{Pr(B)} \times Pr(B) = Pr(A \cap B) = \frac{Pr(A \cap B)}{Pr(A)} \times Pr(A)$$

分母が約分できるから
この等式は当たり前ですね!



そして, 条件付確率

$$Pr(A|B) = \frac{Pr(A \cap B)}{Pr(B)}, \quad Pr(B|A) = \frac{Pr(A \cap B)}{Pr(A)}$$

という表現を思い出せば

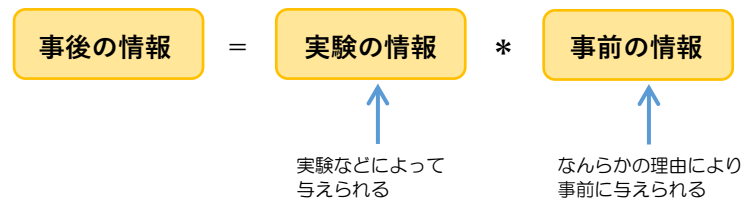
$$Pr(A|B) \times Pr(B) = Pr(B|A) \times Pr(A)$$

となるので, ベイズの定理

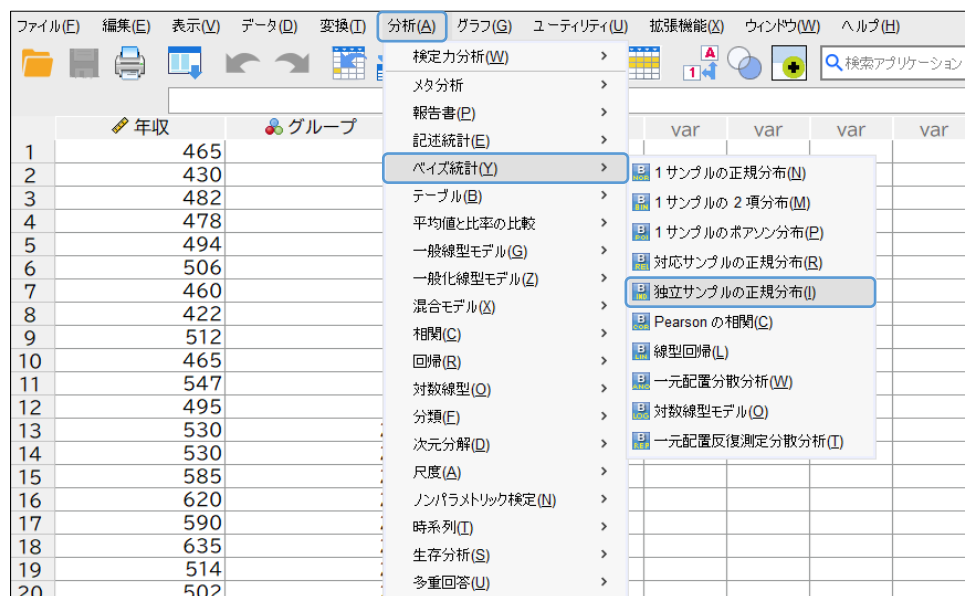
$$Pr(A|B) = \frac{Pr(B|A)}{Pr(B)} \times Pr(A)$$

ができあがります.

したがって、ベイズ統計のイメージは、次のようになります。



SPSS でも、次のようにベイズ統計ができるようになりました。



ここでは、このサブメニューの中の、**独立サンプルの正規分布(I)** を使って、
 “ベイズ統計による 2 つのグループの間の差”
 を調べてみましょう。

次のデータは、事務職のグループと専門職のグループに対しておこなったアンケート調査の結果です。

この 20 人の年収は、次のようになりました。

【データ入力】

	💰 年収	👤 グループ	var
1	465	1	
2	530	1	
3	482	1	
4	478	1	
5	494	1	
6	506	1	
7	460	1	
8	422	1	
9	512	1	
10	465	1	
11	547	1	
12	495	1	
13	530	2	
14	470	2	
15	585	2	
16	550	2	
17	590	2	
18	535	2	
19	514	2	
20	502	2	
21			

値ラベル



	💰 年収	👤 グループ	var
1	465	事務職	
2	530	事務職	
3	482	事務職	
4	478	事務職	
5	494	事務職	
6	506	事務職	
7	460	事務職	
8	422	事務職	
9	512	事務職	
10	465	事務職	
11	547	事務職	
12	495	事務職	
13	530	専門職	
14	470	専門職	
15	585	専門職	
16	550	専門職	
17	590	専門職	
18	535	専門職	
19	514	専門職	
20	502	専門職	
21			

このデータで
やってみます！



【ベイズ統計の手順】

手順① **分析(A)** のメニューから **ベイズ統計(B)** を選択し、
 続いて、サブメニューから **独立サンプルの正規分布(I)** を選びます。

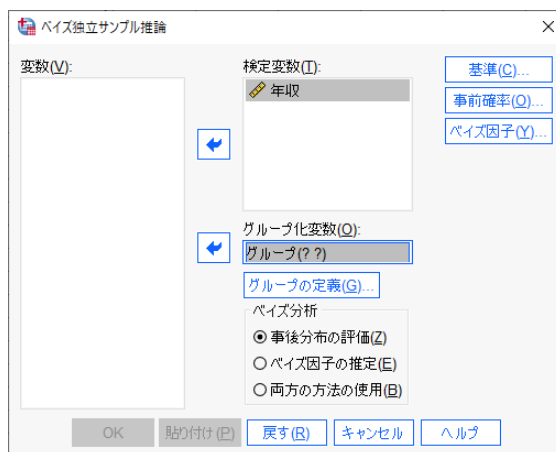
The screenshot shows the SPSS menu bar with '分析(A)' (Analyze) highlighted. The '分析(A)' menu is open, showing various statistical options. The 'ベイズ統計(Y)' (Bayesian Statistics) option is highlighted, and its submenu is also open, showing '独立サンプルの正規分布(I)' (Independent Samples T-Test) as the selected option.

年収	グループ
1	465
2	530
3	482
4	478
5	494
6	506
7	460
8	422
9	512
10	465
11	547
12	495
13	530
14	470
15	585
16	550
17	590
18	535
19	514
20	502
21	
22	
23	
24	
25	
26	
27	
28	
29	
30	
31	



いろいろやって
 みたいけど
 まずはここから！

- 手順② **検定変数(T)** には、**年収** を移動。
グループ化変数(O) には、**グループ** を移動します。



- 手順③ 続いて、**グループの定義(G)** をクリックして、**1** と **2** と入力。
ベイズ分析 のところは、**両方の方法の使用(B)** を選びます。



手順④ **基準(C)** をクリックすると、次の画面になります。
 ここでは、このまま **続行** します。

ベイズ独立サンプル推論: 基準

信頼区間のパーセント %(P): 95

欠損値
☒ 分析ごとに除外(E)
☐ リストごとに除外(W)

適応求積公式
 許容度(I): 0.000001
 最大反復回数(X): 2000

続行 キャンセル ヘルプ

このデータでは
 このまま続行！



手順⑤ 手順 3 の画面に戻ったら、**事前確率(O)** をクリック。
 次のように、**等分散を仮定する(A)** を選んで、**続行**。

ベイズ独立サンプル推論: 事前確率分布

データ分散
☐ 既知の分散(V)
 グループ 1 の分散(1):
 グループ 2 の分散(2):
☒ 等分散を仮定する(A)
☐ 不等分散を仮定する(U)

分散の事前確率
☒ Jeffreys(J)
☐ カイ 2 乗の逆数(I)
 自由度(D):
 尺度パラメータ(S):
☒ 未知の仮定不等分散に拡散事前確率を使用

分散の条件付き平均の事前確率
☒ 拡散(E)
☐ 正規(N)

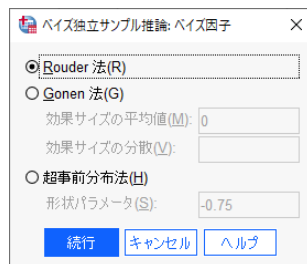
	グループ 1	グループ 2
位置パラメータ		
尺度パラメータ(S)		

続行 キャンセル ヘルプ

2つのグループの
 母分散 σ_1^2 と σ_2^2 は
 等しいと仮定します

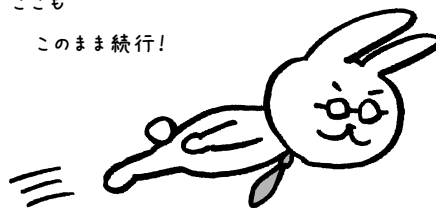


手順⑥ 手順 3 の画面に戻ったら、**ベイズ因子(Y)** をクリック。
 すると、次の画面になります。ここでは、このまま **続行**。



ここも

このまま続行！



手順⑦ 次の画面に戻ってきたら、**OK** ボタンを押します。



【SPSS による出力・その 1】

ベイズ独立

グループ統計量					
	グループ	N	平均値	標準偏差	平均値の標準誤差
年収	= 事務職	12	488.00	33.796	9.756
	= 専門職	8	534.50	40.581	14.348

← ①

ベイズ因子独立サンプルの検定 (方法 = Roudner) ^a						
	平均値の差	プールされた差の標準誤差	ベイズ因子 ^b	t 値	自由度	有意確率 (両側)
年収	46.50	16.699	.217	2.785	18	.012

a. グループ間で等分散を仮定します。

b. ベイズ因子: 帰無仮説 対 対立仮説。

↑

↑

②

③

↑

②

独立サンプル平均値の事後分布評価 ^a					
	事後分布			95% 信用区間	
	最頻値	平均値	分散	下限	上限
年収	46.50	46.50	313.698	11.42	81.58

a. 分散の事前確率: Jeffreys 2。平均の事前確率: Diffuse。

Diffuse とは
無情報事前分布
のことです



効果サイズ

$$d = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s^2}} = -1.271$$

$$r = \sqrt{\frac{t^2}{t^2 + N_1 + N_2 - 2}} = 0.549$$

【出力結果の読み取り方・その1】

←① 2つのグループの基礎統計量です.

標本平均	標本標準偏差	
$\bar{x}_1 = 488.00$	$s_1 = 33.796$	$\frac{s_1}{\sqrt{N_1}} = 9.756$
$\bar{x}_2 = 534.50$	$s_2 = 40.581$	$\frac{s_2}{\sqrt{N_2}} = 14.348$

←② ベイズ因子です.

このベイズ因子は、次の2つのモデルを比較しています.

モデル1 ⇒ 帰無仮説 H_0 : 2つのグループの年収は等しい

モデル2 ⇒ 対立仮説 H_1 : 2つのグループの年収は異なる

$$\text{ベイズ因子} = \frac{\{\text{モデル1}\}}{\{\text{モデル2}\}} = \frac{\{\text{帰無仮説}H_0\}}{\{\text{帰無仮説}H_1\}}$$

ベイズ因子 0.217 は, 1 より小さいので,

“モデル2を支持している”

ことがわかります.



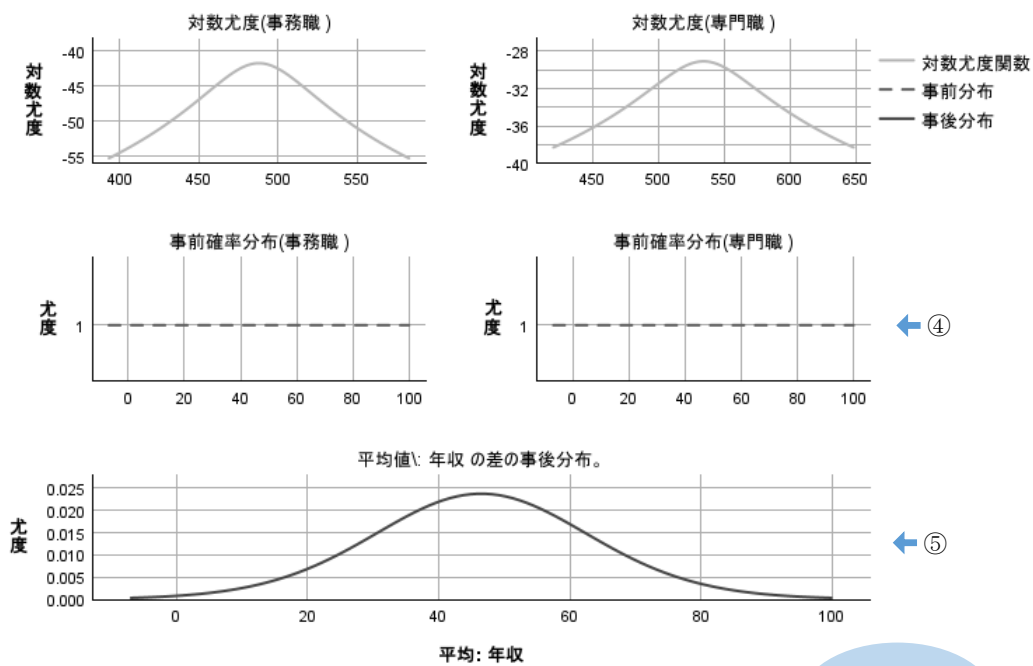
ベイズ因子の評価は
p.12, 13にあります

←③ t 値 2.785 は、等分散を仮定したときの検定統計量です.

有意確率（両側） $0.012 \leq$ 有意水準 0.05

なので、帰無仮説は棄てられます.

【SPSS による出力・その 2】



評価の基準は
いろいろあります

SPSS による ベイズ因子 Bf_{01} の評価

● H_1 を支持する

$0 \sim \frac{1}{100}$	極度に H_1 を支持する
$\frac{1}{100} \sim \frac{1}{30}$	非常に強く H_1 を支持する
$\frac{1}{30} \sim \frac{1}{10}$	強く H_1 を支持する
$\frac{1}{10} \sim \frac{1}{3}$	適度に H_1 を支持する
$\frac{1}{3} \sim 1$	根拠なく H_1 を支持する

【出力結果の読み取り方・その2】

←④ 事前分布を確認しましょう。

無情報事前分布として、一様分布を利用しています。

←⑤ 2つの平均の差の事後分布です。

$$\text{ベイズ因子 } Bf_{01} = \frac{\{\text{帰無仮説 } H_0\}}{\{\text{対立仮説 } H_1\}}$$



これは評価のイメージです

SPSS による ベイズ因子 Bf_{01} の評価

● H_0 を支持する

極度に H_0 を支持する	100 以上
非常に強く H_0 を支持する	30 ~ 100
強く H_0 を支持する	10 ~ 30
適度に H_0 を支持する	3 ~ 10
根拠なく H_0 を支持する	1 ~ 3