基礎・経済統計　５

確率

０．確率（１）

- 事象　　出来事
- 事象（出来事）がおきる可能性の指標
  - 確実に起きるが１，起きないが０
  - 起きる可能性が確実に起きる場合の何％か？
- 様々な事象を表すためにA，Bなど大文字を使用
  - Aが起きる確率をP（A）で表す

０．確率（２）

- 積事象：どちらもが起きるという出来事　$A \cap B$
  - 積事象の確率を同時確率ともいう
- 排反事象
  - AかBのどちらかしか起きないという関係
  - AとBが一緒に起きるということはない
- 和事象：どれかが起きるという出来事　$A \cup B$
  - これはAとBのどちらかしか起きないという意味ではない
  - もし，AとBが排反事象なら，$P(A \cup B) = P(A) + P(B)$

０．確率（３）

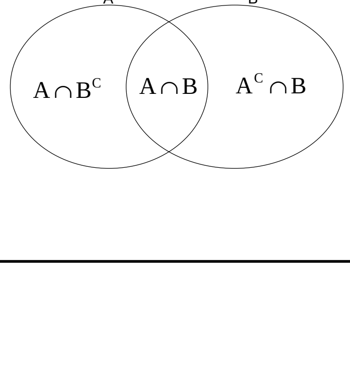
- 余事象
  - 「ある出来事が起きない」という出来事
  - Aの余事象 A^Cで表す
  - $P(A^C) = 1 - P(A)$
    - AとA^Cは排反事象（いずれかしか起こらない）
      $P(A^C \cup A) = P(A) + P(A^C)$
    - さらに，AかA^Cのどちらかが必ず起きる
      $P(A^C \cup A) = 1$
    - $P(A) + P(A^C) = 1$　　$P(A^C) = 1 - P(A)$

０．確率（４）

- $P(A \cup B)$ の計算
  - AかBのどれかが起きる確率（ただし，両方起きてもよい）
  - $A \cup B$ の中身
    - $A \cap B^C$，$A^C \cap B$，$A \cap B$ （これらは排反事象）
    - $P(A \cup B) = P(A \cap B^C) + P(A^C \cap B) + P(A \cap B)$
  - A，Bの中身
    - $A \cap B, A \cap B^c$　　$P(A) = P(A \cap B) + P(A \cap B^C)$
    - $A \cap B, A^C \cap B$　　$P(B) = P(A \cap B) + P(A^C \cap B)$

０．確率（５）

- ベン図

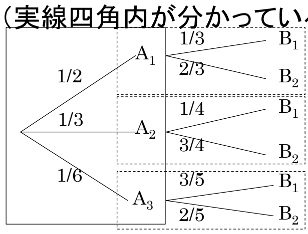


０．確率（６）

- $P(A \cup B) = P(A \cap B^C) + P(A^C \cap B) + P(A \cap B)$ に
  $P(A) = P(A \cap B) + P(A \cap B^C)$ から得られた
  $P(A \cap B^C) = P(A) - P(A \cap B)$ と
  $P(B) = P(A \cap B) + P(A^C \cap B)$ から得られた
  $P(A^C \cap B) = P(B) - P(A \cap B)$ とを代入
- ＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿を得る．

１．条件付き確率
１．１　条件付き確率の概念（１）

- 確率については高校レベルの知識を仮定
  – 本学の入試に出るレベル
- ２段構えの確率＝条件付き確率
  （実線四角内が分かっている情報）



$P(B_1|A_1)$＝1/3
$P(B_2|A_1)$＝2/3
$P(B_1|A_2)$＝1/4
$P(B_2|A_2)$＝3/4
$P(B_1|A_3)$＝3/5

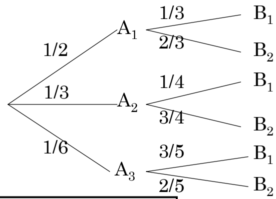
１．１　条件付き確率の概念（２）

- 同じB_1が起きる確率でも，A_1，A_2，A_3のどれが起きたかによって確率が違ってくる．
  – 例
    - 得点圏打率
      – 野球における2塁，3塁に走者がいる場合の打率で，その打者の勝負強さを示す指標．普通，ランナーがいない場合の打率と違ってくる．
    - トランプ
      – すでに，A（エース）が出ている場合と，出ていない場合で，Aが出る確率は違う（教科書p．65例2．9）
      – 確率系ゲーム＝トランプ，麻雀などにおいてはこれを常に意識する必要がある．
    - 交通事故
      – 教科書p.67例2．11やH社のPに乗っている人の事故率が異常に高いなどで，これがリスク細分型自動車保険の基礎になっている

１．２　条件付き確率と同時確率

- 条件付き確率から同時確率を求める

	B_1	B_2
A_1	1/2×1/3 =1/6	1/2×2/3 =1/3
A_2	1/3×1/4 =1/12	1/3×3/4 =1/4
A_3	1/6×3/5 =1/10	1/6×2/5 =1/15

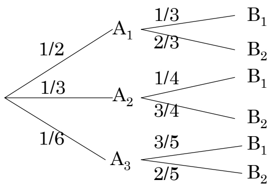


- $P(B_1$かつ$A_1) = P(B_1 \cap A_1) = P(B_1|A_1) \times P(A_1)$
- 一般的：$P(X \cap Y) = P(X|Y) \times P(Y) = P(Y|X) \times P(X)$

１．３　条件付き確率と普通の確率

- 条件付き確率から普通の確率を求める

	B_1	B_2
A_1	1/2×1/3 =1/6	1/2×2/3 =1/3
A_2	1/3×1/4 =1/12	1/3×3/4 =1/4
A_3	1/6×3/5 =1/10	1/6×2/5 =1/15
	$P(B_1)$= 1/6+1/12 +1/10=7/20	$P(B_2)$= 1/3+1/4+ 1/15=13/20



一般式

$$
\begin{aligned}
&P(B_1) \\
&= P(B_1 \cap A_1) + P(B_1 \cap A_2) + P(B_1 \cap A_3) \\
&= P(B_1 | A_1)P(A_1) + P(B_1 | A_2)P(A_2) + P(B_1 | A_3)P(A_3)
\end{aligned}
$$

１．４　同時確率から条件確率へ

- 一般法則

$P(X|Y)=P(X \cap Y)/P(Y)$
$P(Y|X)=P(X \cap Y)/P(X)$

理由：$P(X \cap Y)=P(X|Y) \times P(Y)=P(Y|X) \times P(X)$

- 練習
  - 以下の結合確率の分割表から条件付き確率を求めてみよう

	B_1	B_2
A_1	1/4	1/5
A_2	1/6	1/7
A_3	1/8	97/840

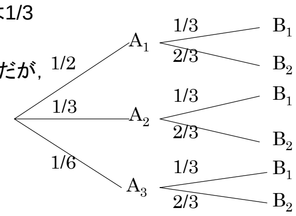
13

２．独立
２．独立性の定義

- 独立
  - 右図の様な場合
  - A_1, A_2, A_3のどれがでようがB_1の起きる確率は1/3
  - 一般的には

$P(X|Y)=P(X)$, 同じことだが，
$P(X \cap Y)$
$=P(X|Y) \times P(Y)$
$= P(X) \times P(Y)$

が成立するとき



14

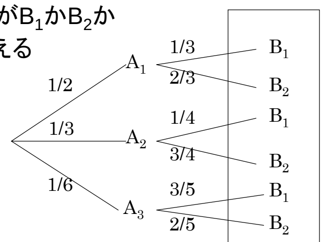
２．２　独立な例

- 得点圏打率の場合
  - 得点圏に走者がいようがいまいが同一の打率である＝バッティングマシーンのような精神的にタフな打者
- 太陽黒点と経済活動
  - 今やsunspot(太陽黒点）均衡というものが経済学の概念ではあるが，たぶん太陽黒点と経済活動は関係ないであろう
  - パルサーB1957+20のX線放出量と経済活動
- 保険の場合
  - いくらリスク細分型保険でも飼っているペットの種類・有無で保険料を変えないだろう．それは，事故確率とこれが独立であるから

15

３．ベイズの定理
３．１　問題設定

- $P(A_1|B_1)$，$P(A_1|B_2)$には意味があるだろうか？
  - 分かっている情報がB_1かB_2かしかない場合を考える
  - B_1が起きていると分かっていてA_1,A_2,A_3のどれが起きているかを確率的に推測する



16

３．２　現実の例

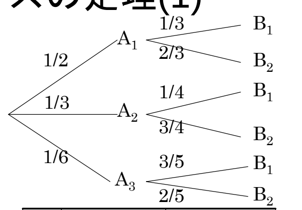
- 医療の例
  - 症状がBにあたり，Aが原因となる病因．医者は，様々な症状によって，確率の高いいくつかの病因に対する対処
- 経済学の例
  - 情報が隠されている場合のインセンティブの設計
    - 情報の経済学
- ビジネスの例
  - Bが注文した品目，Aが顧客の嗜好
    - 本の注文サイトアマゾンが利用しているらしい

17

３．３　ベイズの定理(1)

- 考え方
  - 確率樹形図（ツリー）から分割表を作成
  - 「同時確率から条件確率へ」の式を適用
- 実例
- $P(A_1|B_1)$

$=P(A_1 \cap B_1)/P(B_1)$
$=(1/6) \div (7/20)$
$=10/21$



	B_1	B_2
A_1	1/6	1/3
A_2	1/12	1/4
A_3	1/10	1/15

$P(B_1)=7/20$　$P(B_2)=13/20$

18

3.3 ベイズの定理（2）

- 一般式

$$
\begin{aligned}
&P(A_i \mid B_k) \\
&\quad = \frac{P(A_i \cap B_k)}{P(B_k)} \\
&\quad = \frac{P(B_k \mid A_i)P(A_i)}{P(B_k \mid A_1)P(A_1) + P(B_k \mid A_2)P(A_2) + P(B_k \mid A_3)P(A_3)}
\end{aligned}
$$

19

星座と事故（1）

- 徳島県警曰く「魚座は事故を起こしやすい」
  - 本当か，検証する
  - 右図は徳島県警HPより転載

http://www.police.tokushima.tokushima.jp/11toukei/seiza_kotujiko.html

星座	事故件数	構成率
牡羊座	76	7.2
牡牛座	84	8.0
双子座	88	8.3
蟹座	83	7.9
獅子座	86	8.1
乙女座	75	7.1
天秤座	94	8.9
蠍座	73	6.9
射手座	89	8.4
山羊座	102	9.7
水瓶座	97	9.2
魚座	109	10.3
合計	1,056	100.0

20

星座と事故（2）

- ベイズの法則の応用

$\mathrm{A}_1 = $ 事故を起こす　$\mathrm{A}_2 = $ 事故を起こさない

$\mathrm{B}_i = $ 星座i生まれである

$$
\begin{aligned}
&\mathrm{P}(\text{事故を起こす} \mid \text{i星座生まれ}) \\
&\quad = \mathrm{P}(\mathrm{A}_1 \mid \mathrm{B}_i) = \frac{\mathrm{P}(\mathrm{A}_1 \cap \mathrm{B}_i)}{\mathrm{P}(\mathrm{B}_i)} \\
&\quad = \frac{\mathrm{P}(\mathrm{B}_i \mid \mathrm{A}_1)\mathrm{P}(\mathrm{A}_1)}{\mathrm{P}(\mathrm{B}_i)} \propto \frac{\mathrm{P}(\text{i星座生まれ} \mid \text{事故を起こす})}{\mathrm{P}(\text{i星座生まれ})}
\end{aligned}
$$

事故率係数とここでは呼ぶことにする

21

星座と事故（3）

- つまり，徳島県警謹製のデータを各星座別の出生率で割って，各星座で比較すれば「魚座は本当に事故りやすい」かどうかがわかる
  - 星座別出生データはないので，月別出生数データを用いる（厚生労働省人口動態調査）

	1月	2月	3月	4月	5月	6月	7月	8月	9月	10月	11月	12月
	山羊座	水瓶座	魚座	牡羊座	牡牛座	双子座	蟹座	獅子座	乙女座	天秤座	蠍座	射手座
出生比率（推定）	9.9	8.5	8.7	8.3	7.9	7.5	8.4	8.5	8.4	8.2	7.8	7.9
事故比率	9.7	9.2	10.3	7.2	8.0	8.3	7.9	8.1	7.1	8.9	6.9	8.4
事故率係数	0.7	0.9	1.0	1.0	1.0	0.9	1.1	0.8	1.0	1.2	1.2	1.3

22

星座と事故（4）

- 結果から見ると魚座が事故を起こしやすいと言うことはない
- むしろ10，11，12月生まれが事故を起こしやすく，1，8月生まれが事故を起こしにくいと言うことになるが．．．
  - 出生数の合計は必ずしも現存ドライバーの出生月別分布と一致しない．
  - 徳島県の事故データのサンプル数が少なすぎる
  - 星座と月を無理矢理対応づけている
- さらなる検討が必要

23