7. More on independence of events (Kelly 2.4)

7.1: Independence of multiple events

(i) Pairwise independence (reminder)

Recall E and F are independent if $P(E \cap F) \;=\; P(E) \times P(F)$.

Then $P(E \mid F) \;\equiv\; P(E \cap F)/P(F) \;=\; P(E)$ and $P(F \mid E) \;\equiv\; P(E \cap F)/P(E) \;=\; P(F)$.

Recall that if E and F are independent, so are E and F^c, E^c and F,and E^c and F^c.

(ii) Joint independence

E_1, E_2, ..., E_n are (jointly) independent if for every subset $E_{r_1}, E_{r_2}, ...$ with $r_1 < r_2 < .... \leq n$

$$P(E_{r_1} \cap E_{r_2} \cap .... \cap E_{r_k}) \;=\; P(E_{r_1}) \times P(E_{r_2}) \times .... \times P(E_{r_k}).$$

7.2 Pairwise and joint independence: examples

(i) Pairwise independence without joint independence: example.

Two independent rolls of a fair die. D_1 is first throw gives odd number.

D_2 is second throw gives odd number. D_3 is sum of two throws is odd number.

$P(D_1) \;=\; P(D_2) \;=\; P(D_3) \;=\; 1/2$. $\qquad P(D_1 \cap D_2) \;=\; P(D_1 \cap D_3) \;=\; P(D_2 \cap D_3) \;=\; 1/4$.

But $P(D_1 \cap D_2 \cap D_3) \;=\; 0$, not 1/8. These three events are pairwise independent but NOT jointly independent.

(ii) The three-way independence, without pairwise, is clearly also possible.

Let F, G, I be Swiss adults fluent in French, German and Italian.

Suppose $P(F) \;=\; P(G) \;=\; P(I) \;=\; 1/2$, and $P(F \cap G \cap I) \;=\; P(F) \times P(G) \times P(I) \;=\; 1/8$.

But this does not determine $P(F \cap G)$ etc.

For example, we could have as shown. Then

$P(F \cap G) \;=\; 3/8, \;\; P(F \cap I) \;=\; 1/8, P(G \cap I) = 1/8$.

	I			I^c		
	F	F^c		F	F^c	
G	1/8	0		1/4	1/8	1/2
G^c	0	3/8		1/8	0	1/2
	1/8	3/8		3/8	1/8	

7.3 Repeated sampling with and without replacement

Urn with m balls, k blue balls and $m-k$ white balls

(i) With replacement; successive draws are independent.

On each draw $P(\text{blue}) \;=\; k/m \;=\; p$. Let B_i be event i th ball is blue.

$P(B_1) = p$, $P(B_2) = p$, $P(B_2 \mid B_1) = p$, $P(B_1 \mid B_2) = p$, $P(B_1 \cap B_2) \;=\; p^2$.

Probability of x blue balls in n draws is $\binom{n}{x} p^x (1-p)^{n-x}$.

(ii) Without replacement: there is negative dependence in colors of balls

$P(B_1) = k/m = p$. $P(B_1 \cap B_2) = k(k-1)/m(m-1)$.

$P(B_2 \mid B_1) \;=\; (k-1)/(m-1) \;<\; k/m$, $P(B_2 \mid B_1^c) = k/(m-1) > k/m$.

$P(B_2) \;=\; k(k-1)/m(m-1) + k(m-k)/m(m-1) \;=\; k/m$.

Probability of x blue balls in n draws is $\binom{k}{x}\binom{m-k}{n-x}/\binom{m}{n}$ (hypergeometric).

Lecture 8: More Conditional probability

8.1 Conditional probability is a probability Kelly 2.2.13

1. $P(D \mid E) = P(D \cap E)/P(E) \geq 0$. (we assume $P(E) > 0$.)
2. $P(\Omega \mid E) = P(\Omega \cap E)/P(E) = P(E)/P(E) = 1$.
3. Note $(\cup_i D_i) \cap E = \cup_i (D_i \cap E)$. So, for disjoint D_i,

$$P(\cup_i D_i \mid E) = P(\cup_i (D_i \cap E))/P(E) = \sum_i P(D_i \cap E)/P(E) = \sum_i P(D_i \mid E).$$

So conditional probabilities satisfy all the probability laws. For example,

$$P((C \cup D) \mid E) = P(C|E) + P(D|E) - P(C \cap D \mid E)$$
$$D_1 \subset D_2 \Rightarrow P(D_1 \mid E) \leq P(D_2 \mid E)$$

8.2 Updating information

(i) Bayes' Theorem (again: see lecture notes 5.3).

Assume $P(D)$ and $P(H)$ are both > 0. Then, by definition,

$$P(D \mid H)\ P(H) = P(D \cap H) = P(H \mid D)\ P(D) \quad \text{or} \quad P(H \mid D) = P(D \mid H)\ P(H)\ /\ P(D)$$

Note also, from the law of total probability

$$P(D) = P(D \cap H) + P(D \cap H^c) = P(D|H)P(H) + P(D \mid H^c)(1 - P(H))$$

(ii) Mutually exclusive and exhaustive events H_i ("hypotheses", or "states of nature") (Kelly 2.3.8)

Suppose H_i has probability $P(H_i)$ and $\sum_{i=1}^k P(H_i) = 1$.

$$P(H_i \mid D) = P(D \mid H_i)P(H_i)/P(D) \quad \text{and} \quad P(D) = \left(\sum_{j=1}^k P(D \mid H_j)P(H_j)\right)$$

8.3 Polya Urn model: a model for positive dependence

Urn with m balls, k blue balls and $m-k$ white balls.

When we draw a ball, we replace **and put back additional ball of that color**.

(i) Probability nth ball is blue given r blue in first $(n-1)$ draws.

Before this draw, there are $k+r$ blue out of $m+n-1$ total in urn, so probability is $(k+r)/(m+n-1)$.

Note this does not depend what order the first ones were selected in.

Example: $P(B_3 \mid B_1 \cap B_2^c) = P(B_3 \mid B_1^c \cap B_2) = (k+1)/(m+2)$.

(ii) Probability second ball is blue, given first is blue.

$P(B_2 \mid B_1) = (k+1)/(m+1) > k/m$: positive dependence.

(ii) Probability second ball is blue

$P(B_2) = P(B_2 \mid B_1)P(B_1) + P(B_2 \mid B_1^c)P(B_1^c) = (k+1).k/(m+1).m + k.(m-k)/(m+1).m = k/m.$

This is as for hypergeomtric, 7.3 (ii).

Similarly the probability any ball is blue, **not given any other information** is k/m.

This makes sense given order does not matter, 8.3 (i).

Lecture 9: Updating information using conditional probabilities

9.1 Updating information sequentially

(i) The probability of new data

Suppose again that D and E are independent given each H_i: $P(E \mid D) = \sum_{i=1}^{k} P(E \mid H_i)P(H_i \mid D)$.

Example: two coins C_1 and C_2, with probability head 1/4 and 3/4. Choose one coin randomly and toss it twice. What is the probability the second toss is heads given the first is heads?

Solution 1: $P(\text{2 nd. head} \mid \text{first head}) = P(\text{both heads})/P(\text{1 st head}) =$

$$((1/4) \times (1/4) \times (1/2) + (3/4) \times (3/4) \times (1/2))/((1/4) \times (1/2) + (3/4) \times (1/2)) = 5/8.$$

Solution 2: After first head, $P^*(C_1) = P(C_1 \mid \text{head}) = (1/4) \times (1/2)/((1/4) \times (1/2) + (3/4) \times (1/2)) = 1/4.$

So $P^*(C_2) = P(C_2 \mid \text{head}) = (1 - (1/4)) = 3/4.$

$P(\text{heads again}) = P^*(C_1)P(\text{head} \mid C_1) + P^*(C_2)P(\text{head} \mid C_2) = (1/4) \times (1/4) + (3/4) \times (3/4) = 5/8.$

9.2 The general case

Suppose now we have two data events D and E:

$$P(H_i \mid D \cap E) = P(D \cap E \mid H_i)P(H_i)/\left(\sum_{j=1}^{k} P(D \cap E \mid H_j)P(H_j)\right)$$

Also, provided D and E are independent given each H_i, $i = 1, ..., k$:

$$P(H_i \mid D \cap E) = P(E \mid H_i)P(H_i \mid D)/\left(\sum_{j=1}^{k} P(E \mid H_j)P(H_j \mid D)\right)$$

That is, we can first update from $P(H_i)$ to $P(H_i \mid D)$ and then use these probabilities in updating to $P(H_i \mid D \cap E)$. And then so also for the next event, and the next, ...

9.3 Example from 2008 midterm

(a) In a population 25% of people are type *bb*, 50% are type *bg*, and the remaining 25% have grey eyes, *gg*. If *bb* marries *gg*, all kids have brown eyes. If *bg* marries *gg*, kids are indep. 50/50, brown/grey eyes.

Let B_0 be the event Sarah has brown eyes, *bb* is event Sarah is type *bb*, and *bg* is event Sarah is type *bg*.

(b) Sarah marries Paul, who has grey eyes. Their first child has brown eyes: event B_1.

(c) Sarah and Paul's second child also has brown eyes: event B_2.

(d) Sarah and Paul's third child has grey eyes; event G_3.

	bb	bg	gg	$bb\mid\ldots$	$bg\mid\ldots$		bb	bg	$P^*(bb)$	$P^*(bg)$
-	0.25	0.5	0.25	1/4	1/2		0.25	0.5	1/4	1/2
$\cap B_0$	$\times 1 = 0.25$	$\times 1 = 0.5$	0	1/3	2/3		1	1	1/3	2/3
$\cap B_1$	$\times 1 = 0.25$	$\times 0.5 = 0.25$	-	1/2	1/2		1	0.5	1/2	1/2
$\cap B_2$	$\times 1 = 0.25$	$\times 0.5 = 0.125$	-	2/3	1/3		1	0.5	2/3	1/3
$\cap G_3$	$\times 0 = 0$	$\times 0.5$	-	0	1		0	0.5	0	1

The left side shows the cumulative way of looking at the problem (see also solutions to 2008 midterm).

The right side shows the conditional updating view– the second formula above.

Note it is exact same result and almost exact same computations; the idea is that using the conditional updating form we do not need to know what went before, only the current $P^*(bb)$ and $P^*(bg)$.