Ex $A =$ "I have 2 daughters", $B =$ "I have 1 daughter"

What is $P(A \mid B)$?

$\Omega_0 = \{D, S\}$, $\mathcal{F}_0 = \mathcal{P}(\Omega_0)$, $P_0(D) = 1/2$

$\Omega = \Omega_0 \times \Omega_0 = \{(D,D), (D,S), (S,D), (S,S)\}$

$\mathcal{F} = \mathcal{P}(\Omega)$, $P = P_0 \otimes P_0$ (see previous lectures)

(e.g. $A_1 = \{D\}$, $A_2 = \{D, S\}$, $P(A_1 \times A_2) = P(A_1) \cdot P(A_2)$
$= P_0(D) \cdot P_0(D,S) = \frac{1}{2} \cdot 1 = \frac{1}{2}$)

$A, B \in \mathcal{F}$ $\qquad A \subset B$

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)}{P(B)}$$

$$= \frac{P(D,D)}{P((D,D), (D,S), (S,D))} = \frac{1/4}{3/4} = \frac{1}{3}$$

$C =$ "I have 1 daughter born in the morning"

$P(A \mid C)$

Guess first : $P(A \mid C) \overset{>}{\underset{<}{=}} P(A \mid B)$?

Ω_0 doesn't distinguish between am and pm birth.

$\tilde{\Omega}_0 = \{D^A, D^P, S\}$ $\qquad \tilde{\mathcal{F}}_0 = \mathcal{P}(\tilde{\Omega}_0)$

$\tilde{\Omega} = \tilde{\Omega}_0 \times \tilde{\Omega}_0$, $\tilde{\mathcal{F}} = \mathcal{P}(\tilde{\Omega})$

$\tilde{P}_0(D^A) = \tilde{P}_0(D^P) = \frac{1}{4}$, $\tilde{P}_0(S) = \frac{1}{2}$

$\tilde{P} = \tilde{P}_0 \otimes \tilde{P}_0$

$$A \not\subset C$$

$$P(A|C) = \frac{P(A\cap C)}{P(C)}$$

$$= \frac{P((D^A,D^A),(D^A,D^P),(D^P,D^A))}{P((D^A,D^A),(D^A,D^P),(D^A S),(D^P,D^A),(S,D^A))}$$

$$= \frac{\frac{1}{4}\cdot\frac{1}{4}\cdot 3}{\frac{1}{4}\cdot\frac{1}{4}\cdot 3 + \frac{1}{4}\cdot\frac{1}{2}\cdot 2} = \frac{\cancel{\left(\frac{1}{4}\right)^2}}{\cancel{\left(\frac{1}{4}\right)^2}} \frac{3}{3+4} = \frac{3}{7}$$

$\frac{3}{7} > \frac{1}{3}$, so $P(A|C) > P(A|B)$

(conditioned on smaller set, as sketch above)

"Two children paradox"
(not really a paradox ...)

Condition can have the role of providing information that modifies a belief. For example here

$P(A) = 1/4$ (without additional info)

$P(A|B) = 1/3$

$P(A|C) = 3/7$

Other possible conditions, e.g. very simple

$D =$ "I have a son"

$P(A|D) = 0$

Random variables

$(\Omega, \mathcal{F}, P)$ proba space (outcome space, σ-algebra, proba)

Def: $X : \Omega \longrightarrow \mathbb{R}$ is called random variable if
$\{X<a\} \in \mathcal{F}$ for all $a \in \mathbb{R}$ $\quad (*)$

(Recall short form $\{X<a\} = \{\omega \in \Omega \mid X(\omega) < a\}$)

- Then also $\{X \geq a\} \in \mathcal{F}$, $\{X \leq a\} \in \mathcal{F}$, $\{X=a\} \in \mathcal{F}$, $\{X > a\} \in \mathcal{F}$ etc.
- Hence, all corresponding proba are defined; $P(X \geq a)$, $P(X \leq a)$, $P(X=a)$, $P(X>a)$

Independence of RVs?

Carry over definition used for events.
Two RV X and Y are independent if

$$P(X \leq a, Y \leq b) = P(X \leq a) \cdot P(Y \leq b) \quad \forall a, b \in \mathbb{R}$$

means "and"
short for
$\{X \leq a\} \cap \{Y \leq b\}$

Expectation

X RV on proba. space $(\Omega, \mathcal{F}, P)$

Discrete case: X has values in $\{x_1, x_2, \dots\}$

$$E[X] = \sum_{i=1}^{\infty} x_i \cdot P(X = x_i)$$

Continuous case: X has values in $\mathbb{R}$ with a
(not examinable) probability density function f

$$E[X] = \int_{-\infty}^{\infty} x \underbrace{f(x)\,dx}$$

interpretation: this corresponds to "$P(X=x)$" in the discrete case (using $P(X=x)$ as is would not make sense as often 0 for all x)

Properties of the expectation

X, Y RVs on $(\Omega, \mathcal{F}, P)$, $a \in \mathbb{R}$ (constant)

(i) $E[a] = a$

(ii) $E[X+Y] = E[X] + E[Y]$ "additivity"

(iii) $E[aX] = a \cdot E[X]$

(i) & (ii) $\Rightarrow$ $E[X+a] = E[X] + a$

Note that (iii) is not generally true if the constant is replaced by a RV, but it is a property of X and Y.

X and Y are called uncorrelated if

$$E[X\cdot Y] = E[X]\cdot E[Y]$$

Note: X, Y independent $\Rightarrow$ X, Y uncorrelated
but $\not\Leftarrow$

What about $E[f(X)]$?

For example, $f(x) = x^2$:

constant

$$E[(X-E[X])^2] = E[X^2 - 2X\underbrace{E[X]}_{} + \underbrace{E[X]^2}_{}]$$

$$= E[X^2] - 2\cdot E[X]\cdot E[X] + E[X]^2$$

$$= E[X^2] - E[X]^2$$

This is actually called variance of X,
which is not 0 unless X is constant
(to be precise, $P(X = E[X]) = 1$)

On the other hand,

$$(E[X - E[X]])^2 = (E[X] - E[X])^2 = 0$$

So this is an example that, in general it
is <u>not</u> $E[f(X)] = f(E[X])$

Concrete example: Coin toss $\Omega = \{0,1\}$,
$\mathcal{F} = \{\emptyset, \{0\}, \{1\}, \{0,1\}\}$, $P(X=1) = p$
$E[X] = p\cdot 1 + (1-p)\cdot 0 = p$, $E[X^2] = p\cdot 1^2 + (1-p)\cdot 0^2 = p$,
$Var(X) = E[X^2] - E[X]^2 = p - p^2 = p(1-p) > 0$
for all $p \in (0,1)$