| *Discrete Structures* | *Prelim 1— **Solutions*** |
|---|---|
| *CS2800* | *Selected problems from past exams* |

1. *True or false*

   (a) $\{\emptyset\} = \emptyset$

   **Solution** False. $\{\emptyset\}$ has 1 element, while $\emptyset$ has none.

   (b) *Every set is a subset of its power set*

   **Solution** False. Every set is an *element* of its power set.

   (c) *A set of $n$ events are mutually independent if all pairs of events are independent.*

   **Solution** False.

   (d) $\Pr(E \cup F) = \Pr(E \setminus F) + \Pr(F)$

   **Solution** True by Kolmogorov's third axiom (since $E \setminus F$ and $F$ are disjoint and their union is $E \cup F$).

   (e) *For any random variables $X$ and $Y$, $E(X + 2Y) = E(X) + 2E(Y)$.*

   **Solution** **True.** Expectation of the sum is the sum of the expectations, and constants factor out of expectations.

   (f) *If $E(XY) = E(X)E(Y)$ then $X$ and $Y$ are independent.*

   **Solution** **False.** The converse is true: if $X$ and $Y$ are independent, then $E(XY) = E(X)E(Y)$, but the expectations could still be the same even if $X$ and $Y$ are not independent.

   (g) *A proof that starts "Choose an arbitrary $y \in \mathbb{N}$, and let $x = y^2$" is likely to be a proof that $\forall y \in \mathbb{N}, \forall x \in \mathbb{N}, \ldots$.*

   **Solution** **False.** This would only be a proof that $\exists x \in \mathbb{N}$ with some property, not a proof that $\forall x \in \mathbb{N}$ the property holds.

   (h) *If $A$ and $B$ are independent events, then $P(A|B) = P(A \cap B)/P(B)$.*

   **Solution** **True.** This is the definition of $P(A|B)$ and holds even if $A$ and $B$ are not independent.

   (i) *The logical negation of "everybody can fool Mike" is "nobody can fool Mike".*

   **Solution** **False.** The negation would be "somebody can't fool Mike".

   (j) *For any random variable $X$, $E(X^2) = (E(X))^2$*

**Solution** **False.** If this were true, then the variance of any random variable would be zero. Alternatively, choose any random variable with $E(X) = 0$; then $X^2 \geq 0$ so $E(X^2) \geq 0$ and is only 0 if $X$ is always 0.

2. *Suppose we pick a bit string of length 4 at random, all bit strings equally likely. Consider the following events:*

   $E_1$*: the string begins with 1.*

   $E_2$*: the string ends with 1.*

   $E_3$*: the string has exactly two 1's.*

   (a) *Find* $\Pr(E_1)$, $\Pr(E_2)$, *and* $\Pr(E_3)$.

   **Solution** 1/2, 1/2, and 3/8 respectively.

   (b) *Find* $\Pr(E_1 \mid E_3)$.

   **Solution** 1/2

   (c) *Find* $\Pr(E_2 \mid E_1 \cap E_3)$.

   **Solution** 1/3

   (d) *Are* $E_1$ *and* $E_2$ *independent? Justify your answer.*

   **Solution** Yes, $\Pr(E_1 \cap E_2) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = \Pr(E_1) \cdot \Pr(E_2)$

   (e) *Are* $E_2$ *and* $E_3$ *independent? Justify your answer.*

   **Solution** Yes, $\Pr(E_2 \cap E_3) = \frac{3}{16} = \frac{1}{2} \cdot \frac{3}{8} = \Pr(E_2) \cdot \Pr(E_3)$

3. *The following questions refer to successive rolls of a fair six-sided die, where each roll yields an integer* $m$ *in the range* $1 \leq m \leq 6$ *independently with uniform probability.*

   (a) *What is the expected number of times that 2 is rolled in 12 rolls?*

   **Solution** 2

   (b) *What is the expected value of each roll?*

   **Solution** $3\frac{1}{2}$

   (c) *What is the expected sum of four rolls?*

   **Solution** 14

4. *The following questions refer to three independent flips of a fair coin.*

   (a) *Give an example of three events that are mutually independent.*

   **Solution** first flip heads, second flip heads, third flip heads

   (b) *Give an example of three events that are pairwise independent but not mutually independent.*

**Solution** first flip heads, second flip tails, first and second flip the same

(c) *Give an example of three events, no pair of which are independent.*

**Solution** three heads, three tails, two heads and one tail

5. *Briefly and clearly identify the errors in each of the following proofs:*

(a) **Proof that 1 is the largest natural number:** *Let $n$ be the largest natural number. Then $n^2$, being a natural number, is less than or equal to $n$. Therefore $n^2 - n = n(n-1) \leq 0$. Hence $0 \leq n \leq 1$. Therefore $n = 1$.*

**Solution** The error is in the first sentence "Let $n$ be the largest natural number". The proof is only valid *if* there is a largest natural number (which there isn't).

(b) **Proof that 2 = 1:** *Let $a = b$.*

$$\begin{aligned}
\Rightarrow \quad & a^2 = ab \\
\Rightarrow \quad & a^2 - b^2 = ab - b^2 \\
\Rightarrow \quad & (a+b)(a-b) = b(a-b) \\
\Rightarrow \quad & a + b = b
\end{aligned}$$

*Setting $a = b = 1$, we get 2 = 1.*

**Solution** The error comes when we divide both sides by $(a - b)$, which is zero (division by zero is meaningless!). Just because $(a - b)x = (a - b)y$, we cannot conclude that $x = y$.

(c) **Proof that $(a + b)(a - b) = a^2 - b^2$:**

$$\begin{aligned}
\text{To prove:} \quad & (a+b)(a-b) = a^2 - b^2 \\
\Rightarrow \quad & a^2 - ab + ab - b^2 = a^2 - b^2 \\
\Rightarrow \quad & a^2 - b^2 = a^2 - b^2
\end{aligned}$$

*... which is true, hence the result is proved.*

**Solution** Although the claim is actually true, the proof is backwards; it begins by assuming that the claim is true, and then derives a fact that is known to be true.

This is a valid proof that *if* $(a + b)(a - b) = a^2 - b^2$ *then* $a^2 - b^2 = a^2 - b^2$, but this is not a very interesting fact (and is not what was claimed).

6. *If $A$ and $B$ are any two events in a probability space $(S, P)$, prove using (only) Kolmogorov's axioms and basic set theory that $P(A \cup B) \leq P(A) + P(B)$.*

**Solution** $A \cup B = A \cup (B \setminus A)$. Since $A$ and $B \setminus A$ are disjoint, we can apply Kolmogorov's third axiom to conclude that $P(A \cup B) = P(A) + P(B \setminus A)$.

I claim that $P(B \setminus A) \leq P(B)$. This is because $B$ is the union of the disjoint sets $B \setminus A$ and $B \cap A$, and thus $P(B) = P(B \setminus A) + P(B \cap A)$. Since $P(B \cap A) \geq 0$ (by axiom 2), we conclude that $P(B \setminus A) \leq P(B)$.

Combining this with the above statement, we conclude that $P(A \cup B) = P(A) + P(B \setminus A) \leq P(A) + P(B)$.

7. *Let $(\mathcal{S}, \mathcal{P})$ be a probability space. Prove that $\mathcal{P}(\emptyset) = 0$. You may use any results from set theory without proof.*

**Solution** We know that $\mathcal{S} \cup \emptyset = \mathcal{S}$. Moreover, $\mathcal{S} \cap \emptyset = \emptyset$. Thus, by axiom 3, $\mathcal{P}(S) = \mathcal{P}(S) + \mathcal{P}(\emptyset)$. Subtracting $\mathcal{P}(\mathcal{S})$ from both sides, we see that $\mathcal{P}(\emptyset) = 0$ as required.

8. *Let $(\mathcal{S}, \mathcal{P})$ be a probability space, and let $A$ and $B$ be events of $\mathcal{S}$.*

   (a) *Give the definition of "$A$ and $B$ are independent".*

   **Solution** $P(A \cap B) = P(A)P(B)$.

   (b) *Give the definition of $\mathcal{P}(A \mid B)$.*

   **Solution** $P(A \mid B) = P(A \cap B)/P(B)$

   (c) *Assume $\mathcal{P}(B) \neq 0$. Prove that if $A$ and $B$ are independent then $\mathcal{P}(A \mid B) = \mathcal{P}(A)$.*

   **Solution** By definition, $\mathcal{P}(A \mid B) = P(A \cap B)/P(B)$. Since $A$ and $B$ are independent, $P(A \cap B) = P(A)P(B)$. Thus $P(A \mid B) = P(A)P(B)/P(B) = P(A)$ as required.

9. *Sid and Mike go to the ice cream store. Sid likes chocolate ice cream, but he doesn't want to order exactly the same flavour that Mike orders. If Mike does not order chocolate ice cream, Sid will order it 90% of the time. If Mike does order chocolate ice cream, Sid will order it only 30% of the time. Mike chooses first. There are 5 different flavours of ice cream (only one is chocolate), and Mike chooses a flavour completely at random (i.e. equiprobably). If Sid ends up buying chocolate ice cream, what is the probability Mike also ordered chocolate ice cream?*

**Solution** **Version 1** We will consider a sample space with 4 outcomes: Mike and Sid both get chocolate (which I will write $C_{MS}$), Mike gets chocolate and Sid does not ($C_M$), Sid alone gets chocolate ($C_S$) and nobody gets chocolate $C$.

Using the assumptions in the problem, we see that $\mathcal{P}(\{C_{MS}\}) = 0.2 \cdot 0.3$, $\mathcal{P}(\{C_M\}) = 0.2 \cdot 0.7$, $\mathcal{P}(\{C_S\}) = 0.8 \cdot 0.9$ and $\mathcal{P}(\{C\}) = 0.8 \cdot 0.1$.

We wish to compute

$$\begin{aligned}
\mathcal{P}(\text{Mike gets chocolate} \mid \text{Sid gets chocolate}) &= \mathcal{P}(\{C_{MS}, C_M\} \mid \{C_{MS}, C_S\}) \\
&= \mathcal{P}(\{C_{MS}\})/\mathcal{P}(\{C_{MS}, C_S\}) \\
&= 0.2 \cdot 0.3/(0.2 \cdot 0.3 + 0.8 \cdot 0.9) \approx 0.08
\end{aligned}$$

**Version 2** Let $M$ be the event Mike orders chocolate ice cream, and $S$ be the event Sid orders chocolate ice cream. Then

$$\begin{aligned}
P(M) &= 1/5 = 0.2 \\
P(M') &= 1 - 0.2 = 0.8 \\
P(S|M) &= 0.3 \\
P(S|M') &= 0.9
\end{aligned}$$

By Bayes' Theorem, $P(M|S) = P(S|M)P(M)/P(S)$

By the Total Probability Theorem, $P(S) = P(S|M)P(M) + P(S|M')P(M')$.

Plugging in the values, we find

$$P(M|S) = 0.3 \cdot 0.2/(0.3 \cdot 0.2 + 0.9 \cdot 0.8) \approx 0.08$$

10. *Which of these is the correct negation of* $\exists x, \neg\forall y, \neg\exists z, \neg F(x,y,z)$*?*

   (a) $\exists x, \exists y, \exists z, F(x,y,z)$

   (b) $\exists x, \exists y, \exists z, \neg F(x,y,z)$

   (c) $\forall x, \forall y, \forall z, F(x,y,z)$

   (d) $\forall x, \forall y, \forall z, \neg F(x,y,z)$

   **Solution** (c).

$$\begin{aligned}\neg\left(\exists x, \neg\forall y, \neg\exists z, \neg F(x,y,z)\right) &= \forall x, \neg\neg\left(\forall y, \neg\exists z, \neg F(x,y,z)\right)\\ &= \forall x, \forall y, \left(\neg\exists z, \neg F(x,y,z)\right)\\ &= \forall x, \forall y, \forall z, \left(\neg\neg F(x,y,z)\right)\\ &= \forall x, \forall y, \forall z, F(x,y,z)\end{aligned}$$

11. *Consider a disease in which one out of a hundred people has. Given a test that will answer "yes" with probability* $999/1000$ *if a person has the disease and will answer "yes" with probability* $2/1000$ *is the person does not have the disease. If the test says "yes" what is the probability that the person has the disease?*

   **Solution** Let $D$ be the event that the person has the disease, and let $T$ be the event that the test says "yes". Then we are given

$$\begin{aligned}P(T|D) &= .999\\ P(T|\bar{D}) &= .002\\ P(D) &= .01\end{aligned}$$

   By Bayes's rule:

$$\begin{aligned}P(D|T) &= \frac{P(T|D)P(D)}{P(T)}\\ &= \frac{(.999)(.01)}{P(T|D)P(D) + P(T|\bar{D})P(\bar{D})}\\ &= \frac{(.999)(.01)}{(.999)(.01) + (.002)(0.99)} \approx 1/2\end{aligned}$$

12. (a) *Give the definition of variance in terms of expectation.*

   **Solution**

$$\mathrm{Var}(X) = \mathrm{E}((X - \mathrm{E}(X))^2)$$

   I would also accept $\mathrm{Var}(X) = \mathrm{E}(X^2) - \mathrm{E}^2(X)$.

   (b) *Let* $X$ *and* $Y$ *be random variables with* $\mathrm{E}(X) = \mathrm{E}(Y) = 0$*. Prove that* $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$*. Make (and clearly state) additional assumptions if necessary.*

**Solution** We must require that $X$ and $Y$ are independent. Therefore $\mathrm{E}(XY) = 0$. Then we have

$$\begin{aligned}
\mathrm{Var}(X+Y) &= \mathrm{E}\left((X+Y-\mathrm{E}(X+Y))^2\right) && \text{by definition} \\
&= \mathrm{E}\left((X+Y)^2\right) && \text{since } \mathrm{E}(X+Y) = \mathrm{E}(X) + \mathrm{E}(Y) = 0 \\
&= \mathrm{E}(X^2 + 2XY + Y^2) && \text{arithmetic} \\
&= \mathrm{E}(X^2) + 2\mathrm{E}(XY) + \mathrm{E}(Y^2) && \text{expectation of the sum is sum of expectations} \\
&= \mathrm{E}(X^2) + \mathrm{E}(Y^2) && \text{since } \mathrm{E}(XY) = 0 \text{ (independence)} \\
&= \mathrm{Var}(X) + \mathrm{Var}(Y) && \text{by definition and fact that } E(X) = E(Y) = 0
\end{aligned}$$

13. (a) *The average human height is 5 feet and 4 inches, and the variance is 2 squared inches. How large a sample must I take so that my estimate of the average will (with 90% probability) be correct to within a half inch?*

**Solution** Let $H$ denote the "height" random variable, and let $E$ be the estimated average.

The law of large numbers states that $\mathcal{P}(|E - E(H)| \geq \epsilon) \leq \sigma^2(H)/n\epsilon^2$. Plugging in a half inch for epsilon and solving for $n$, we see that if $n \geq 2/(0.25 \cdot 0.1)$ that the probability of being incorrect is no larger than 0.1.

(b) *A certain high school is divided into two teams: 35% of the students are "beliebers", and the remaining 65% are "directioners".*

*90% of the songs on a belieber's playlist will be Justin Bieber songs, while the other 10% will be by One Direction. Directioners are a bit more broad-minded: 80% of their songs will be One Direction songs, while the remaining 20% will be Justin Bieber songs.*

*A student is selected at random, and a random song is selected from their playlist. It turns out to be "Baby" by Justin Bieber. What is the probability that the student was a directioner?*

**Solution** Let $D$ be the event that the student is a directioner, and let $B$ be the event that the student is a belieber. Let $S_D$ be the event that the selected song was a One Direction song, and $S_B$ be the event that the selected song was a Justin Bieber song. We are given

- $P(B) = 35\%$
- $P(D) = 65\%$
- $P(S_D|D) = 80\%$
- $P(S_D|B) = 10\%$
- $P(S_B|D) = 20\%$
- $P(S_B|B) = 90\%$

We wish to find $P(D|S_B)$. The correct answer is given by Bayes' rule:

$$P(D|S_B) = \frac{P(S_B|D)P(D)}{P(S_B)} = \frac{P(S_B|D)P(D)}{P(S_B|D)P(D) + P(S_B|B)P(B)} = \frac{.2 \cdot .65}{.2 \cdot .65 + .9 \cdot .35}$$

14. *Let $A$ and $B$ be two* independent *events in a probability space $(S, P)$, that is,*

$$P(A \cap B) = P(A)P(B)$$

*Prove that $A$ and $B'$ (the complement of $B$, i.e. $S \setminus B$) are also independent. That is, prove that*

$$P(A \cap B') = P(A)P(B')$$

*You may use any results proved in class without proof.* **Hint:** *Observe that* $A = (A \cap B) \cup (A \cap B')$.

**Solution** Following the hint, we write $A$ as $(A \cap B) \cup (A \cap B')$. Since $A \cap B$ and $A \cap B'$ are disjoint, we can apply Axiom 3 to obtain

$$\begin{aligned}
& P(A) = P((A \cap B) \cup (A \cap B')) \\
\implies \quad & P(A) = P(A \cap B) + P(A \cap B') && \text{(Axiom 3)} \\
\implies \quad & P(A) = P(A)P(B) + P(A \cap B') && \text{(independence)} \\
\implies \quad & P(A) - P(A)P(B) = P(A \cap B') \\
\implies \quad & P(A)(1 - P(B)) = P(A \cap B') \\
\implies \quad & P(A)P(B') = P(A \cap B') && \text{(P(B') = 1 - P(B), proved in class)}
\end{aligned}$$

Hence $A$ and $B'$ are independent.

15. *Grumpy Cat and Happy Cat are engaged in a perennial war. The war is played out over several individual battles, also known as "lectures". In a particular semester, the instructors are Mike and Sid. Mike teaches 25 lectures, and Sid teaches 15 lectures. When Mike lectures, Happy Cat wins 70% of the time; and when Sid lectures, Grumpy Cat wins 90% of the time. You missed a lecture that semester, and asked your friend to fill you in. The friend mentioned in passing that Grumpy Cat won that lecture. What is the probability that Sid taught the lecture?*

**Solution** Use Bayes' Theorem for the following events:

- $M$: Mike taught lecture.
- $S$: Sid taught lecture.
- $G$: Grumpy Cat won lecture.
- $H$: Happy Cat won lecture.

$$P(S|G) = \frac{P(G|S)P(S)}{P(G|S)P(S) + P(G|M)P(M)} = \frac{\frac{9}{10} \times \frac{15}{40}}{\frac{9}{10} \times \frac{15}{40} + \frac{3}{10} \times \frac{25}{40}} = \frac{9}{14}$$

16. *Let* $S = \{a, b\}$, *and the function* $P$ *given by the following table:*

| $A$ | $P(A)$ |
|---|---|
| $\emptyset$ | 0 |
| $\{a\}$ | 1/4 |
| $\{b\}$ | 3/4 |
| $\{a, b\}$ | 1 |

(a) *What are the events of* $S$*?*

**Solution** We helpfully listed them for you when defining $P$! They are all possible subsets of $S$, i.e. $\emptyset$, $\{a\}$, $\{b\}$, and $\{a, b\}$.

(b) *Prove that* $(S, P)$ *is a probability space.*

**Solution** We're given that $S$ is a set, and we've just shown that the domain of $P$ covers all possible subsets of that set. The *only* correct way to wrap up the argument is to show that all three of Kolmogorov's axioms hold. Let's take them one by one.

**Does Axiom 1 hold?** The probability of every event is non-negative – we listed all the events above, and their probabilities are explicitly given in the table. Hence Axiom 1 holds.

**Does Axiom 2 hold?** $P(S) = P(\{a,b\}) = 1$, again from the table. So Axiom 2 holds.

**Does Axiom 3 hold?** This is a little tedious but feasible since $S$ is so small. Let's first consider non-empty subsets:

- $P(\{a\} \cup \{b\}) = P(\{a,b\}) = 1$, $P(\{a\}) + P(\{b\}) = 1/4 + 3/4 = 1$. Yep.

...and that's it! For this set, this is the only possible way to break it into disjoint non-empty subsets (note that for the third axiom, the order of the subsets does not matter).

What happens if we include empty subsets? Well, let's take any number of instances of the empty set and add them to any other countable set of disjoint non-empty subsets $A_1, A_2, \ldots$:

$$\begin{aligned} P(A_1 \cup A_2 \cup \cdots \cup \emptyset \cup \emptyset \cup \ldots) &= P(A_1 \cup A_2 \cup \ldots) \qquad \text{(by definition, } X \cup \emptyset = X\text{)} \\ &= P(A_1 \cup A_2 \cup \ldots) + 0 + 0 + \ldots \\ &= P(A_1) + P(A_2) + \cdots + 0 + 0 + \ldots \qquad \text{(we just proved this)} \\ &= P(A_1) + P(A_2) + \cdots + P(\emptyset) + P(\emptyset) + \ldots \qquad \text{(from the table)} \end{aligned}$$

This covers all possible countable sets of disjoint subsets of $A$. So we have proved Axiom 3 holds. (We were a little lenient when grading this part – e.g. we gave you credit even if you didn't handle repetitions of the empty set explicitly.)

Since all three of Kolmogorov's axioms hold, $(S, P)$ is a probability space.

*(c) Are $\{a\}$ and $\{b\}$ independent? Explain.*

**Solution** The only way this could be true is if $P(\{a\} \cap \{b\}) = P(\{a\}) \cdot P(\{b\})$. But this is not the case: $P(\{a\} \cap \{b\}) = P(\emptyset) = 0$, and $P(\{a\}) \cdot P(\{b\}) = 1/4 \times 3/4 \neq 0$. So the events are not independent. Note that this is the only possible correct answer, derived from the definition of independence.

17. *Let $E$ be a herd of 100 elephants. The herd contains 10 adult males, 60 adult females and 30 babies. It is known[1] that the adult elephants have an average surface area of 17m$^2$, and the babies have an average surface area of 4m$^2$. A biologist, unaware of these statistics, picks an elephant uniformly at random from $E$ and measures its surface area (after temporarily and painlessly tranquilizing it). If the measured surface area is represented as a random variable, what are its (a) domain, (b) codomain, and (c) expectation (show your calculations)?*

(Note: *There are many correct answers for (a) and (b). Pick any one.)*

**Solution**

(a) Domain: $E$, or $\{\text{adult}, \text{baby}\}$, or $\{\text{adult male}, \text{adult female}, \text{baby}\}$, or other reasonable variation.

(b) Codomain: $\mathbb{R}$, or $\mathbb{R}^+$, or $\{17, 4\}$, or $\{17\text{m}^2, 4\text{m}^2\}$, or other reasonable variation.

(c) Expectation: the numerical answer (in m$^2$) is $17 \times 0.7 + 4 \times 0.3 = 13.1$. There are many ways to calculate this, depending on how you choose your domain. E.g. let's choose the domain $E$, call the random variable $A$, and for notational convenience (without loss of generality) assume the adult

[1] K. P. Sreekumar and G. Nirmalan, "Estimation of the total surface area in Indian elephants (Elephas maximus indicus)", Veterinary Research Communications, 1990;14(1):5-17.

elephants are numbered 1 to 70, and the young 'uns 71-100. Then you might write:

$$\begin{aligned}
\text{Expectation of } A &= \sum_{i=1}^{100} A(elephant_i)P(elephant_i) \\
&= \sum_{i=1}^{70} A_{adult}P(elephant_i) + \sum_{i=71}^{100} A_{baby}P(elephant_i) \\
&= \sum_{i=1}^{70} 17 \cdot \frac{1}{100} + \sum_{i=71}^{100} 4 \cdot \frac{1}{100} \\
&= 17 \cdot \frac{70}{100} + 4 \cdot \frac{30}{100}
\end{aligned}$$

18. *How would you (humanely) measure the surface area of an elephant?*

**Solution** There were many excellent answers, but a favorite was: "very carefully".

19. *In an infamous criminal case, a mother was accused of murdering her two infant sons. A well-known statistician testified that the chance that* both *deaths were natural was infinitesimal. He proposed the following calculation:*

$$P(D_1 \cap D_2 \mid I) = P(D_1 \mid I) \cdot P(D_2 \mid I)$$

*where $D_1$ and $D_2$ are the events that the two children respectively died, and $I$ is the event that the mother is innocent.*

*Since natural infant death is rare in the family's demographic, both probabilities on the right hand side are tiny: about 1/8543. Plugging in the values, we obtain $P(D_1 \cap D_2 \mid I) \approx$ 1/73,000,000. Based on this, the mother was found guilty and imprisoned.*

*Four years later, the ruling was overturned on grounds of faulty statistics. There are* **two** *significant errors in the reasoning above. Briefly and clearly identify both.*

**Solution** The first error is that the two deaths are presumed independent (conditioned on innocence). This is unjustified: genetic factors etc. can increase the chances of multiple deaths in the same family.

The second error is trickier but more damning: the conditional of interest is $P(I \mid D_1 \cap D_2)$, not $P(D_1 \cap D_2 \mid I)$. If this is written out via Bayes' Theorem and the Theorem of Total Probability,

$$P(I \mid D_1 \cap D_2) = \frac{P(D_1 \cap D_2 \mid I)P(I)}{P(D_1 \cap D_2 \mid I)P(I) + P(D_1 \cap D_2 \mid I')P(I')},$$

the even smaller chance that a mother would actually murder both her sons makes $P(I \mid D_1 \cap D_2) > 1/2$ (e.g. try plugging in $P(I') = 1$ in a billion. This type of error is known as the "prosecutor's fallacy".

Some of you suggested that the error was in not accounting for someone else being the murderer. In other words, you were saying the problem lies in the statement "Since natural infant death is rare in the family's demographic, both probabilities on the right hand side are tiny." It is true that natural infant death is only one of the ways in which the deaths could happen while the mother was innocent — an axe murderer could also be stalking the neighborhood. While I hope axe murderers aren't common enough that this significantly changes the probability, it is true that there is something slightly fishy here, so we gave you credit. (We hadn't intended this to be the error, btw. We realized later that our wording was slightly off.)

The error is *not* that 1 in 73 million is still not 0. If we required *100 percent* certainty (if such a thing exists), no legal case would ever get settled.

This is an actual case, btw. See `http://en.wikipedia.org/wiki/Sally_Clark`.

20. *Suppose that a coin has probability .6 of landing heads. You flip it 100 times. The coin flips are all mutually independent.*

   (a) *What is the expected number of heads?*

   (b) *What upper bound does Markov's Theorem give for the probability that the number of heads is at least 80?*

   (c) *What is the variance of the number of heads for a* single *toss? Calculate the variance using either of the equivalent definitions of variance.*

   (d) *What is the variance of the number of heads for 100 tosses? You may use the fact that if $X_1, \ldots, X_n$ are mutually independent, then $\mathrm{Var}(\sum X_i) = \sum \mathrm{Var}(X_i)$; you don't need to prove this.*

   (e) *What upper bound does Chebyshev's Theorem give for the probability that the number of heads is either less than 40 or greater than 80?*

**Solution** (a) Let $X_i$ be the outcome of the $i$th coin toss; $X_i = 1$ if the $i$th coin toss lands heads and 0 otherwise. The total number of heads is $Y = X_1 + \cdots + X_{100}$. We are interested in $E(Y)$. By linearity, $E(Y) = E(X_1) + \cdots + E(X_{100}) = 100(.6) = 60$.

(b) Markov's Theorem says that $\Pr(Y \geq 80) \leq E(Y)/80 = 60/80 = 3/4$.

(c) $X_i$ is a Bernoulli variable with $p = .6$, so as shown in class, its variance is $p(1-p) = .24$. You can also compute this directly, since $X_i^2 = X$, so $E(X_i^2) = .6$ and $E(X_i)^2 = .36$, so $Var(X) = E(X_i^2) - E(X_i)K^2 = .6 - .36 = .24$.

(d) $Var(Y) = Var(X_1) + \cdots + Var(X_{100})$. Since $Var(X_i) = .24$ for $i = 1, \ldots, 100$, $Var(Y) = 100 \times .24 = 24$.

(e) By Chebyshev's Theorem. $\Pr(|Y - E(Y)| \geq 20) \leq \mathrm{Var}(Y)/400$. Since $E(Y) = 60$, $\Pr(Y \geq 80 \cup Y \leq 40) \leq 24/400 = .06$.

21. *Let $E$ and $H$ be events in a probability space. We say that $E$ is* evidence in favor of $H$ *if* $\Pr(H|E) > \Pr(H)$. *Similarly, $E$ is* evidence against $H$ *if* $\Pr(H|E) < \Pr(H)$. *Show that if $E$ is evidence in favor of $H$ then $\overline{E}$ is evidence against $H$. (Assume that* $0 < \Pr(E) < 1$.*)*

**Solution** Suppose that $E$ is evidence in favor of $H$. Thus, $\Pr(H \mid E) = \Pr(H \cap E)/\Pr(E) > \Pr(H)$, so $\Pr(H \cap E) > \Pr(H)\Pr(E)$. Now $\Pr(H) = \Pr(H \cap E) + \Pr(H \cap \overline{E})$, so $\Pr(H \cap E) = \Pr(H) - \Pr(H \cap \overline{E})$. It follows that $\Pr(H) - \Pr(H \cap \overline{E}) > \Pr(H)(1 - \Pr(\overline{E}) = \Pr(H) - \Pr(H)\Pr(\overline{E})$. Subtracting $\Pr(H)$ from both sides gives $-\Pr(H \cap \overline{E}) > -\Pr(H)\Pr(\overline{E})$, or equivalently $\Pr(H \cap \overline{E}) < \Pr(H)\Pr(\overline{E})$. Diving both sides by $\Pr(\overline{E})$, we get that $\Pr(H \mid \overline{E}) = \Pr(H \cap \overline{E})/\Pr(\overline{E}) < \Pr(H)$; that is, $\overline{E}$ is evidence against $H$.

22. *Complete the following sentences carefully:*

   (a) Pr *is is a probability measure on a sample space $S$ if* $\Pr : 2^S \to [0,1]$ *satisfying ...*

   (b) *An event in $S$ is ...*

   (c) *A real-valued random variable on $S$ is ...*

   (d) *Two random variables $X$ and $Y$ on $S$ are independent if ...*

**Solution**

   (a) Pr is is a probability measure on a sample space $S$ if $\Pr : Pow(S) \to [0,1]$ such that $\Pr(S) = 1$ and $\Pr(A \cup B) = \Pr(A) + \Pr(B)$ if $A \cap B = \emptyset$.

   (b) An event in $S$ is a subset of $S$.

   (c) A real-valued random variable on $S$ is a function from $S$ to the real numbers.

(d) Two random variables $X$ and $Y$ on $S$ are independent if for all values $x$ in the range of $X$ and all values $y$ in the range of $Y$, the events $X = x$ and $Y = y$ are independent.

23. *A busy student must complete 3 problem sets before doing laundry. Each problem set requires 1 day with probability* $1/4$ *and 2 days with probability* $1/2$ *and 3 days with probability* $1/4$*. The problem sets are done in sequence (the first one is done, then the second, then the third.) Let* $B$ *be the number of days a busy student delays laundry. (For example, if the first problem set requires 1 day and the second and third problem sets each require 2 days, then the student delays for* $B = 5$ *days.) What is the expected value of* $B$*? Explain carefully how you got your answer.*

**Solution** Let $D_i$ be the number of days that problem set $i$ takes, for $i = 1, 2, 3$. Then $B = D_1 + D_2 + D_3$. By the linearity of expectation, $E(B) = E(D_1) + E(D_2) + E(D_3)$. We are told that $E(D_1) = E(D_2) = E(D_3) = 1 \times 1/4 + 2 \times 1/2 + 3 \times 1/4 = 2$. Thus, $E(B) = 6$.

24. *One penny in a barrel of 100 pennies is double-headed. All the rest have heads on one side and tails on the other. All the coins other than the double-headed coin are fair (i.e., each side is equally likely to come up). A friend chooses a penny from the barrel at random, tosses it 10 times, and gets 10 heads.*

(a) *Carefully describe an appropriate sample space for this problem. How many elements does it have?*

(b) *What is the probability that your friend's coin is double-headed?*

**Solution** (a) A reasonable sample space can be described in terms of a probability tree. First you choose a coin, and then toss it 10 times. If you choose a double-headed coin, then there is only one possible sequence of coin tosses; if you choose a fair coin, there are $2^{10}$ possible sequences of coin tosses. Thus, this sample space has $2^{10} + 1$ elements.

(b) Let $DH$ be the event that your friend chose the double headed coin (so $\overline{DH}$ is the event that he chose a fair coin) and let $TH$ be the event that your friend tosses 10 heads in a row. We are interested in $\Pr(DH \mid TH)$. By Bayes' Rule,

$$\Pr(DH \mid TH) = \frac{\Pr(TH \mid DH)\Pr(DH)}{\Pr(TH \mid DH)\Pr(DH) + \Pr(TH \mid \overline{DH})\Pr(\overline{DH})}.$$

We are told that $\Pr(DH) = \frac{1}{1{,}000{,}000}$. $\Pr(TH \mid DH) = 1$ (if you have a double-headed coin, you definitely get 10 heads in a row!), and $\Pr(TH \mid \overline{DH}) = 1/2^{10}$. Thus,

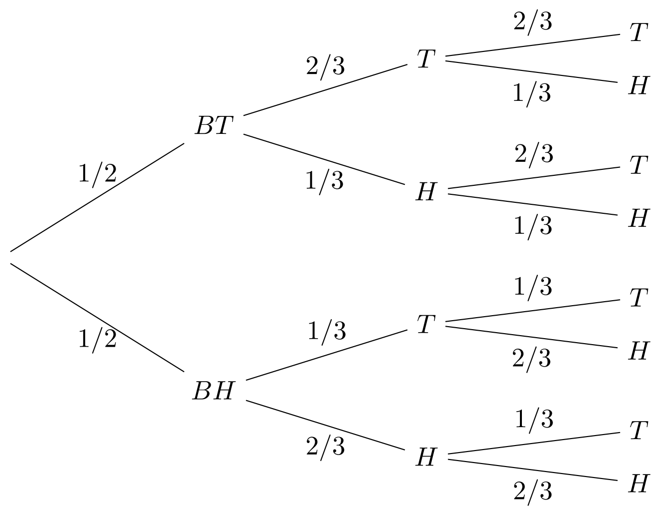
$$\Pr(DH \mid TH) = \frac{\frac{1}{1{,}000{,}000}}{\frac{1}{1{,}000{,}000} + \frac{1}{2^{10}}\frac{999{,}999}{1{,}000{,}000}} = \frac{2^{10} + 999,9999}{2^{10} \times 1,000,000}.$$

(Since $2^{10} \approx 1,0000$, this probability is about 1/1000.)

25. *Suppose that you have two coins. One is biased towards heads: the probability of heads is* $2/3$ *and the probability of tails is* $1/3$*; the other is biased towards tails: the probability of heads is* $1/3$ *and the probability of tails is* $2/3$*. You choose one of the two at random (with equal probability), and then toss it twice.*

(a) *Draw the probability tree.*

**Solution** Here is the probability tree (where $BH$ denotes "biased towards heads" and $BT$ denotes "biased towards tails):




(b) *What's the probability of getting heads the first time you toss the coin? Give a one sentence explanation.*

**Solution** Let $H1$ be the event of getting heads the first time that the coin is tossed. By the law of total probability, $\Pr(H1) = \Pr(BH) \times \Pr(H1 \mid BH) + \Pr(BT) \times \Pr(H1 \mid BT) = 1/2(1/3) + 1/2(2/3) = 1/2$.

(c) *Are the two coin tosses independent?*

**Solution** Let $H2$ be the event of getting heads the second time that the ocin is tossed. By symmetry, $\Pr(H2) = 1/2$. But the probability of getting heads both times is

$$\Pr(H1 \cap H2) = \Pr(BH) \times \Pr(H1 \cap H2 \mid BH) + \Pr(BT) \times \Pr(H1 \cap H2 \mid BT) = 1/2(1/9) + 1/2(4/9) = 5/18.$$

Since $\Pr(H1 \cap H2) \neq \Pr(H1) \times \Pr(H2)$, the coin tosses are not independent. (intuitively, this is because getting heads on the first toss tells you it's more likely that the coin is biased towards heads, which makes it more likely that you'll get heads on the second toss.)

(d) *Suppose you win* \$3 *if the* first *coin lands heads and lose* \$5 *if it lands tails. What are your expected earnings?*

**Solution** Let $X$ be the random variable describing my earnings. $X$ can take the values 3 or $-5$. It takes the value 3 with probability $(1/2)(2/3) + (1/2)(1/3) = 1/2$, and takes $-5$ with probability 1/2 as well. Thus $E(X) = 3/2 - 5/2 = -1$.

(e) *What is the variance of your earnings?*

**Solution** Let $Y = (X - E(X))^2$. If $X = 3$ then $Y = 16$ and if $X = -5$ then $Y = 16$. Thus, $Var(X) = E(Y) = 16$.

26. *A simplified form of Bayes's rule is given by the following expression:*

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

*Prove this identity.*

**Solution** By defintion, $\Pr(A|B) = \Pr(A \cap B)/\Pr(B)$. Thus $\Pr(A \cap B) = \Pr(B)\Pr(A|B)$. Similarly, $\Pr(B \cap A) = \Pr(A)\Pr(B|A)$. Thus $\Pr(A)\Pr(B|A) = \Pr(B)\Pr(A|B)$. Dividing both sides by $\Pr(A)$ gives the formula.

27. *(This is a variant of a problem due to Lewis Carroll, who wrote "Alice in Wonderland".) A bag has a white ball in it. A second ball is put into the bag, which is white with probability* $2/3$ *and black with probability* $1/3$ *(so with probability* $2/3$*, there are two white balls in the bag and with probability* $1/3$ *there is a white ball and a black ball). Now a ball is chosen from the bag at random.*

   (a) *Carefully describe a sample space for this problem.*

   **Solution** The sample space consists of two balls in the bag and the ball that is chosen. Let $W$ be the white ball that was originally in the bag, let $W'$ be the second white ball, and let $B$ be the black ball. Thus, the sample space has four elements:

   - $(WW', W)$ (ball $W'$ is put in the bag, $W$ is chosen)
   - $(WW', W')$ (ball $W'$ is put in the bag, $W$ is chosen)
   - $(WB, W)$
   - $(WB, B)$

   (It's OK to omit the $W$ that is in the bag all along.)

   (b) *What is the probability of each element of this sample space.*

   **Solution** The first two elements in the space above ($(WW', W)$ and $(WW', W')$), where the white ball $W'$ is put in, each have probability $1/3$; the second two elements each have probability $1/6$.

   (c) *If the ball chosen is white, what is the probability that the second ball is white?*

   **Solution** The ball chosen is white in the first three elements of the sample space. The second ball is white in the first two. Thus, the conditional probability is $(\frac{1}{3} + \frac{1}{3})/(\frac{1}{3} + \frac{1}{3} + \frac{1/6}{)} = \frac{4}{5}$.

28. *Let $S$ be a sample space. Give a formal mathematical definition of the following:*

   (a) *Probability measure on $S$*

   **Solution** A function $\Pr : 2^S \to \mathbb{R}$ satisfying (1) for all $E \subseteq S$, $0 \leq \Pr(E) \leq 1$, (2) $\Pr(\emptyset) = 0$ and $\Pr(S) = 1$, and (3) if $E_1 \cap E_2 = \emptyset$ then $\Pr(E_1 \cup E_2) = \Pr(E_1) + \Pr(E_2)$.

   (b) *Event of $S$*

   **Solution** A subset $E \subseteq S$.

   (c) *Random variable on $S$*

   **Solution** A function $X : S \to \mathbb{R}$.

   (d) *Independent random variables of $S$*

   **Solution** Random variables $X$ and $Y$ such that for all $x$ and $y$, the events $X = x$ and $Y = y$ are independent.

   (e) $\Pr(A \mid B)$*, where $A$ and $B$ are events.*

**Solution** $\Pr(A \cap B)/\Pr(B)$.

29. *Roughly 1% of Cornell's undergraduate students take CS 2800. 95% of the CS 2800 students know the correct definition of "injective", while only 20% of students who didn't take 2800 know the definition. While walking down the hall, you overhear one undergraduate student saying to another "... since $f$ is injective, we know that if $x \neq y$ then $f(x) \neq f(y)$, so ...". What is the probability that the student has taken CS 2800? (Assume that you heard a randomly chosen Cornell undergrad.)*

**Solution** Let $C$ be the event that a randomly selected student took CS 2800, and let $I$ be the event that a randomly selected student knows what injective means.

We are given $P(C) = 0.01$, $P(I|C) = 0.95$ and $P(I|\bar{C}) = 0.20$. By Bayes's rule:

$$P(C|I) = \frac{P(I|C)P(C)}{P(I|C)P(C) + P(I|\bar{C})P(\bar{C})} = \frac{.95 \cdot 0.01}{0.95 \cdot 0.01 + 0.20 \cdot 0.99}$$

30. *You toss a blue coin that lands heads 1/2 of the time, and a red coin that lands heads 1/3 of the time. You toss the blue coin 5 times and the red coin 3 times. Suppose you get \$2 every time the blue coin lands heads, and you lose \$1 every time the red coin lands heads. (You get nothing if either coin lands tails.)*

   (a) *Write down a sample space that describes this situation. How many elements does it have?*

   **Solution** There are many possible answers, but the most obvious sample space consists of all possible combinations of results for the 8 flips. For example, $HHHHHTTT$ and $HTHTHHTH$ are all outcomes. There are $2^8$ outcomes in this space.

   (b) *Let $E$ be the event that the third toss of the red coin lands tails. What is* $\Pr(E)$*?*

   **Solution** 2/3

   (c) *Let $F$ be the event that the red coin lands heads at least twice. What is* $\Pr(F)$*?*

   **Solution** The probability that it lands heads 3 times is $(1/3)^3$. The probability that it lands heads twice is the sum of the probabilities of the outcomes where it lands twice. There are $C(3, 2)$ such outcomes, and the probability of each of them is $(1/3)^2(2/3)$. Since landing 3 heads and landing 2 heads are mutually exclusive, the total probability is thus $(1/3)^3 + C(3, 2)(1/3)^2(2/3) = 7/27$.

   (d) *Are $E$ and $F$ independent?*

   **Solution** Note that $E \cap F$ is the event that the third toss of the red coin is tails and the red coin lands at least heads twice. Since there are only three coin tosses (ignoring the blue coin), the events in $E \cap F$ are $HHT$. THis happens with probability 2/27.

   From above, we know $\Pr(E) = 2/3$ and $\Pr(F) = 7/27$. Since $\Pr(E)\Pr(F) = 7/81 \neq 2/27 = \Pr(E \cap F)$, $E$ and $F$ are not independent.

   (e) *Define a random variable that describes how much money you win. Find its expected value.*

   **Solution** Let $X(s)$ be two times the number of blue heads in $s$ minus one times the number of red heads in $s$. If $X_i$ is one if coin $i$ comes up heads and 0 otherwise, then $X = 2(X_1 + \cdots + X_5) - (X_6 + X_7 + X_8)$. Note that $E(X_1) = E(X_2) = \cdots = E(X_5) = 1/2$ and $E(X_6) = E(X_7) = E(X_8) = 1/3$. Thus by linearity of expectation we have $E(X) = 5 - 1 = 4$.

31. *Suppose $X$ is the constant random variable $c$ (that is $X(s) = c$ for all $s$ in the sample space). Show that (a) $E(X) = c$ and (b) $\mathrm{Var}(X) = 0$.*

**Solution** (a) $E(X) = \sum_{s \in S} X(s) \Pr(\{s\}) = c \sum_{s \in S} \Pr(\{s\}) = c$.

(b) First note that $X^2(s) = c^2$ for all $s$, so that $X^2$ is also a constant. Thus $E(X^2) = c^2$ Then we have $\text{Var}(X) = E(X^2) - E(X)^2 = E(c^2) - c^2 = c^2 - c^2 = 0$ as required.

Alternatively, we know $\text{Var}(X) = E((X - E(X))^2) = E((X - c)^2) = E(0) = 0$.

32. *The average height of an adult American is about 5.5 feet, and the standard deviation is about 0.2 feet. You wish to build a door that guarantees that 90% of American adults can enter without ducking. Using Chebychev's inequality, how tall must the door be? (*Hint: *If $h \leq E(H) + d$ then $|h - E(H)| \leq d$. Thus $\Pr(H \geq E(H) + d) \leq \Pr(|H - E(H)| \geq d)$).*

**Solution** If $H \geq E(H) + d$ then $|H - E(H)| \geq d$. Thus $\Pr(H \geq E(H) + d) \leq \Pr(|H - E(H)| \geq d)$.

By Chebychev's inequality, $\Pr(|H - E(H)| \geq d) \leq \text{Var}(H)/d^2$. We wish to find $d$ such that $\text{Var}(H)/d^2 \leq 0.1$. Note that $\text{Var}(H) = 0.2^2 = 0.04$. Thus our requirement is satisfied if $d^2 \geq 0.04/0.1 = 0.4$, so $d \geq \sqrt{.4}$.

Thus the door must be $E(h) + d$ feet (which is about 6.1 feet).

I will not fit in your door ☹ (but most of you will ☺!)