Discrete Structures	*Prelim 2 sample questions*— ***Solutions***
CS2800	*Selected problems from past exams*

1. *In a permutation of the set* $\{1, 2, \ldots, n\}$*, a pair* i, j *is* out of order *if* $i < j$ *but* i *occurs after* j *in the permutation. In a random permutation of the set* $\{1, 2, \ldots, n\}$ *with all permutations equally likely, what is the expected number of pairs that are out of order?*

   Solution $\binom{n}{2}/2$

2. *Consider the following assertion.*

   > *Assuming everyone has three initials, there are at least 6 people in California (population 38 million) with the same initials and the same birthday (the same day of the year, but not necessarily the same year).*

   Describe a simple calculation you could do to verify this assertion.

   Solution Using the extended pigeonhole principle, check whether $\left\lceil \frac{38{,}000{,}000}{366 \cdot 26^3} \right\rceil \geq 6$.

3. *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 $\frac{1}{2}$, $\frac{1}{2}$, $\frac{3}{8}$

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

   Solution $\frac{1}{2}$

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

   Solution $\frac{1}{3}$

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

4. *Give an expression describing the number of different ways the following things can happen. No credit will be given for just the value, even if correct.*

(a) *During your pregnancy, you decided on a list of 23 girls' first names and 16 boys' first names, as well as a list of 11 gender-neutral middle names. To your surprise, you had quintuplets, two boys and three girls. Now you must select a first and a middle name for each child from the lists. The names must all be different.*

Solution $P(16, 2) \cdot P(23, 3) \cdot P(11, 5)$ or equivalent, e.g. $(16!/14!)(23!/20!)(11!/6!)$, $16 \cdot 15 \cdot 23 \cdot 22 \cdot 21 \cdot 11 \cdot 10 \cdot 9 \cdot 8 \cdot 7$

(b) *A professor teaching discrete structures is making up a final exam. He has a stash of 24 questions on probability, 16 questions on combinatorics, and 10 questions on logic. He wishes to put five questions on each topic on the exam.*

Solution $\binom{24}{5}\binom{16}{5}\binom{10}{5}$ or equivalent

(c) *The very same professor wants to assign points to the 15 problems so that each problem is worth at least 5 points and the total number of points is 100.*

Solution $\binom{25+15-1}{25}$ or $\binom{25+15-1}{15-1}$ or equivalent

(d) *There are 30 graders to grade the final exam, and the professor would like to assign two graders to each of the 15 problems.*

Solution $\binom{30}{2\ 2\ 2\ 2\ 2\ 2\ 2\ 2\ 2\ 2\ 2\ 2\ 2\ 2\ 2}$ or $\frac{30!}{2^{15}}$ or $\binom{30}{2}\binom{28}{2}\binom{26}{2}\cdots\binom{2}{2}$

5. *Give an expression describing the probability of the following events. Evaluate the expression if it is easy to do so.*

(a) *A fair coin is flipped 100 times giving exactly 50 heads.*

Solution $\binom{100}{50}2^{-100}$

(b) *A fair coin is flipped 100 times giving at most 50 heads.*

Solution $\sum_{n=0}^{50}\binom{100}{n}2^{-100}$ or $\frac{1}{2}+\frac{1}{2}\binom{100}{50}2^{-100}$. (Our original solutions said 1/2, which was incorrect. Student answers of 1/2 received full credit.)

(c) *A roll of two fair dice yields a sum of 7.*

Solution 1/6

(d) *The conditional probability that the last card dealt in a 5-card poker hand yields a straight (five cards in sequence, irrespective of suit) given that the first four cards dealt were* $5\spadesuit$, $6\heartsuit$, $7\diamondsuit$, $8\clubsuit$.

Solution $1/6$

(e) *A controversial bill before the Senate is supported by 50 of 55 Democrats who will vote for the bill and opposed by 40 of 45 Republicans who will vote against the bill. Of the remaining 10 undecided Senators, each Democrat will vote in favor with probability 1/2 and each Republican against with probability 9/10, independent of the other votes. The bill requires 51 votes to pass. What is the probability that it passes?*

Solution $1 - .9^5 \cdot .5^5$

6. *Give an expression describing the number of different ways the following things can happen. Your expression may involve binomial coefficients, multinomial coefficients, Stirling numbers of the second kind, factorial expressions, r-permutations, or whatever else you need, but do not evaluate the expression. No credit will be given for just the value, even if correct.*

(a) *You must choose a password consisting of 6, 7, or 8 letters from the 26-letter English alphabet* $\{a, b, \ldots, z\}$.

Solution $26^6 + 26^7 + 26^8$

(b) *In a poker game, you are dealt a* full house, *a five-card hand containing three of a kind and a pair of another kind; for example, three kings and two sixes.*

Solution $13 \cdot \binom{4}{3} \cdot 12 \cdot \binom{4}{2}$

(c) *Your team is in the championship game of a soccer tournament. The score is tied at full time and the winner will be decided by penalty kicks. As coach, you must choose a sequence of five different players out of 11 to take the kicks.*

Solution $P(11, 5) = 11!/(11-5)! = 11 \cdot 10 \cdot 9 \cdot 8 \cdot 7$

(d) *You have ten rings, all with different gemstones. You wish to bequeath them to your five children so that each child inherits two of the rings.*

Solution $\binom{10}{2\ 2\ 2\ 2\ 2} = \dfrac{10!}{2!\,2!\,2!\,2!\,2!}$

(e) *You have \$400 to donate to charity, which you would like to distribute among your five favorite charities so that each receives an integral number of dollars.*

Solution $\binom{400+5-1}{400}$

(f) *The same as (e), but you also wish to ensure that each of the five charities receives at least \$20.*

Solution $\binom{300+5-1}{300}$

7. *Give the* formal mathematical definition *of each of the following terms. Your definitions should be valid for a finite or countably infinite sample space S.*

   (a) *Probability distribution on S.*

   Solution A function $\Pr : S \to \mathbb{R}$ such that for all $s \in S$, $\Pr(s) \geq 0$ and $\sum_{s \in S} \Pr(s) = 1$.

   (b) *Event.*

   Solution A subset of S.

   (c) *Probability of an event E, given a probability distribution* Pr *on S.*

   Solution $\Pr(E) = \sum_{s \in E} \Pr(s)$.

   (d) *Conditional probability of E given F.*

   Solution $\Pr(E \mid F) = \Pr(E \cap F) / \Pr(F)$, undefined if $\Pr(F) = 0$.

   (e) *Real-valued random variable X on S.*

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

   (f) *Expectation of a random variable X.*

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

   (g) *Variance of a random variable X.*

   Solution $V(X) = E((X - E(X))^2) = E(X^2) - E(X)^2$.

8. *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

(d) *What is the expected number of rolls before seeing the first 6?*

Solution 6

9. *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

10. *Give an expression describing the probability of the following events. You do not need to evaluate the expression.*

   (a) *You obtain at least three consecutive heads in four flips of a fair coin.*

   Solution 3/16

   (b) *In a poker game, you are dealt a five-card hand containing four aces.*

   Solution $48/\binom{52}{5}$

   (c) *The conditional probability that the first flip of two flips of a fair coin comes up heads, given that at least one of them does.*

   Solution 2/3

   (d) *A 13-card bridge hand contains all cards of the same suit.*

   Solution $4/\binom{52}{13}$

   (e) *A randomly chosen number between 0 and 99 (inclusive) is divisible by 4.*

   Solution 1/4

11. *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) **(1 point)** *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) **(3 points)** *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 $A_1, A_2, \ldots$:

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

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) **(1 point)** *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.

12. **(3 points)** *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.)*

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

(b) *Codomain: $\mathbb{R}$, or $\mathbb{R}^+$, or $\{17, 4\}$, or $\{17m^2, 4m^2\}$, 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 elephants are numbered 1 to 70, and the young 'uns 71-100. Then you might write:*

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

[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.

13. **(2 points)** *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`.

14. **(0 points)** *How would you (humanely) measure the surface area of an elephant?*

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