Discrete Structures	Prelim 1
CS2800	Selected problems from past exams

1. True or false
   (a) $\{\emptyset\} = \emptyset$
   (b) Every set is a subset of its power set
   (c) A set of n events are mutually independent if all pairs of events are independent.
   (d) $\Pr(E \cup F) = \Pr(E \setminus F) + \Pr(F)$
   (e) For any random variables X and Y, $E(X + 2Y) = E(X) + 2E(Y)$.
   (f) If $E(XY) = E(X)E(Y)$ then X and Y are 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 $\forall y \in \mathbb{N}, \forall x \in \mathbb{N}, \ldots$.
   (h) If A and B are independent events, then $P(A|B) = P(A \cap B)/P(B)$.
   (i) The logical negation of "everybody can fool Mike" is "nobody can fool Mike".
   (j) For any random variable X, $E(X^2) = (E(X))^2$

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 $\Pr(E_3)$.
   (b) Find $\Pr(E_1 \mid E_3)$.
   (c) Find $\Pr(E_2 \mid E_1 \cap E_3)$.
   (d) Are E_1 and E_2 independent? Justify your answer.
   (e) Are E_2 and E_3 independent? Justify your answer.

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?
   (b) What is the expected value of each roll?
   (c) What is the expected sum of four rolls?

4. The following questions refer to three independent flips of a fair coin.
   (a) Give an example of three events that are mutually independent.
   (b) Give an example of three events that are pairwise independent but not mutually independent.
   (c) Give an example of three events, no pair of which are independent.

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 $n = 1$.

(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 $2 = 1$.

(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.

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 $P(A \cup B) \leq P(A) + P(B)$.

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.

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".
   (b) Give the definition of $\mathcal{P}(A \mid B)$.
   (c) Assume $\mathcal{P}(B) \neq 0$. Prove that if A and B are independent then $\mathcal{P}(A \mid B) = \mathcal{P}(A)$.

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?

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

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?

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

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?

(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?

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 $A = (A \cap B) \cup (A \cap B')$.

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?

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?

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

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

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.)

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

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.

[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.

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.

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?

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 $0 < \Pr(E) < 1$.)

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 ...

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.

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?

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.

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

    (c) Are the two coin tosses independent?

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

(e) What is the variance of your earnings?

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.

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.

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

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

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

    (a) Probability measure on S

    (b) Event of S

    (c) Random variable on S

    (d) Independent random variables of S

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

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.)

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?

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

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

    (d) Are E and F independent?

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

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) $\text{Var}(X) = 0$.

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 $\Pr(H \geq E(H) + d) \leq \Pr(|H - E(H)| \geq d)$).