Review sheet

Preparation for Midterm 1

Elementary manipulations of sets

Homework: I-2.

Exercise A. A customer buys ice cream at a parlour. Their choice is random, and the possibilities are as follows: the ice cream can be one or two scoops, and each scoop can be strawberry, chocolate or vanilla.

Write S and D for the event that the customer chooses one or two scoops ('single', 'double'), and B, C and V for the events that they choose at least one scoop of (straw)berry, chocolate or vanilla.

1. Write the event "The customer chooses only vanilla" as a formula using only the events described above. $B^c \cap C^c$

2. Same question for the event "The customer chooses at most one strawberry scoop".
$(S \cap B) \cup (D \cap B^c) \cup (D \cap B \cap V) \cup (D \cap B \cap C)$ or $(D \cap C^c \cap V^c)^c$

Exercise B. A baking show has three contestants initially: Andrei, Betty and Charlotte. During the first round, two contestants will pass the challenge, and one will lose and leave the game. During the second one, only one of the remaining two will pass and win the super baking crown.

Write A_1 (resp. B_1, C_1) the event that Andrei (resp. Betty, Charlotte) passes the first round, and A_2, B_2, C_2 the events that they pass the second round. Give two examples of events that are (non-trivially) included,[1] and two examples of events that are disjoint.
$A_2 \subset A_1$ $A_2 \subset B_2^c$ $A_2 \cap B_2 = \emptyset$ $A_1^c \cap B_1^c = \emptyset$

Exercise C. We choose an American household at random. Denote by P, C, D, O the events that the household has at least one pet, at least one cat, at least one dog, and at least one pet that is not a cat nor a dog ('other').

1. Write the event "The household has at least one pet that is not a cat" as a formula using only the events described above. $D \cup O$

2. Same question for the event "The household has at least one pet, but no cat". $P \cap C^c$

3. Same question for the event "The household has one or many dogs, but nothing else". $D \cap C^c \cap O^c$

[1] Two events E and F are included one in the other if $E \subset F$ or $F \subset E$, and the inclusion is trivial if $E = F$.

Axioms of probability

Textbook: 1.1, 1.2, 1.3, 1.5.
Homework: I-3.

Exercise D. Two events A and B are defined on some probability space. Which of the following probabilities are compatible with the axioms of probability? (There may be several good answers.)

1. $\mathbb{P}(A) = 50\%$, $\mathbb{P}(B) = 75\%$, $\mathbb{P}(A \cap B) = 20\%$. Incompatible. $(\mathbb{P}(A \cup B) > 1)$
2. $\mathbb{P}(A \cup B) = 70\%$, $\mathbb{P}(A) = 30\%$, $\mathbb{P}(B) = 30\%$. Incompatible. $(\mathbb{P}(A \cap B) < 0)$
3. $\mathbb{P}(A) = 50\%$, $\mathbb{P}(B) = 50\%$, $\mathbb{P}(A \cap B) = 0\%$. Compatible.
4. $\mathbb{P}(A \cup B) = 75\%$, $\mathbb{P}(A) = 50\%$, $\mathbb{P}(A \cap B) = 50\%$. Compatible.

Exercise E. You play They Love Me, They Love Me Not with a daisy: as you say the first 'They love me', you take off the first petal, then 'They Love Me Not' as you take off the second, and so on (so that they love you if the flower has 3 petals, for instance).

Daisies have between 8 to 12 petals, with probabilities[2]

$$\mathbb{P}(8 \text{ petals}) = \mathbb{P}(12 \text{ petals}) = \frac{1}{12}, \quad \mathbb{P}(9 \text{ petals}) = \mathbb{P}(11 \text{ petals}) = \frac{1}{4}, \quad \mathbb{P}(10 \text{ petals}) = p.$$

1. What should p be so that the probability space satisfies the axioms of probability? $p = \frac{1}{3}$
2. What it the probability that your sweetheart loves you? $\frac{1}{4} + \frac{1}{4} = \frac{1}{2}$.

Exercise F. The probability that a randomly chosen American household has at least a cat is $\mathbb{P}(C) = 25\%$, at least a dog is $\mathbb{P}(D) = 38\%$, a TV is $\mathbb{P}(TV) = 96\%$, a computer is $\mathbb{P}(PC) = 90\%$.

1. What are the possible values for $\mathbb{P}(C \cup D)$? $\mathbb{P}(C \cup D) \in [38\%, 63\%]$
2. What are the possible values for $\mathbb{P}(TV \cap PC)$? $\mathbb{P}(TV \cap PC) \in [86\%, 90\%]$

Counting

Textbook: 1.6, 1.7, 1.8, 1.9, 1.10, 1.11, 1.12, 1.13, 1.14, 1.15, 1.16.
Homework: II-2, II-3.

Exercise G. Dumblebee, headmaster of the paleontology school Joedarts, organises each year the Fourhouse Tournament. He chooses 4 students at random out of the 400 in the school (100 for each house) to compete in various dinosaur-related tasks.

1. What is the probability that at least two students are chosen in the house Goldfirebird?

Dumblebee is notoriously biased towards Goldfirebird, and he will choose a set of four students uniformly out of all those that contain at least two students from this house.

[2]Not biologically accurate.

$$2 \quad P_1 = \frac{\binom{100}{2} \cdot \binom{300}{2} + \binom{100}{3} \cdot \underset{\overset{\|}{300}}{\binom{300}{1}} + \binom{100}{4}}{\binom{400}{4}}$$

$$\frac{\binom{100}{4}}{\binom{100}{2}\binom{300}{2}+300\cdot\binom{100}{3}+\binom{100}{4}} = \frac{\binom{100}{4}/\binom{400}{4}}{P_1}$$

2. What is the probability that all chosen students are from house Goldfirebird?

3. What is the probability that Dumblebee's favourite student Morty Parker (of course from house Goldfirebird) is chosen?

$$\left[1\cdot\binom{99}{1}\cdot\binom{300}{2}+1\cdot\binom{99}{2}\binom{300}{1}+1\cdot\binom{99}{3}\right] \Big/ \left[\binom{100}{2}\cdot\binom{300}{2}+\binom{100}{3}\cdot 300+\binom{100}{4}\right]$$

Exercise H. It seems like any three-letter word constructed in the form consonant-vowel-consonant is a valid English word. Here, we assume the English alphabet has 26 letters (from A to Z) and 5 vowels: A, E, I, O and U.

1. How many words of this form can we construct? $21^2\cdot 5$

2. You are playing Scrabble against an opponent whose native language is not English, and you are fairly confident that you can convince him that any combination of this form is a genuine word. You have the letters B, H, L, T, V, A, O and you are the first to play, so nothing is on the board. How many such 'words' can you form? $(21)_2\cdot 2$

3. Same question with the letters B, H, T, T, V, A and O. $(21)_2\cdot 2-2$

TOT and TAT

Exercise I. You are dealt 5 cards of a standard deck of 52 cards in order.

1. What is the probability that you get all cards from the same suit? $=4\cdot(13)_5/(52)_5$

2. What is the probability that you get all cards in strict increasing order? $=\binom{13}{5}\cdot 4^5/(52)_5$.

3. What is the probability that you get at most two different ranks?

$$\left[\binom{5}{2}\cdot(13)_2\cdot(4)_2\cdot(4)_3+\binom{5}{1}\cdot(13)_2\cdot 4\cdot(4)_4\right]/(52)_5$$

Exercise J. We throw a die 20 times.

1. What is the probability that exactly five of the throws are 2? $\binom{20}{5}\cdot\frac{1}{6^5}\cdot\left(\frac{5}{6}\right)^{15}$

2. What is the probability to get only even numbers? $\left(\frac{3}{6}\right)^{20}$

3. What is the probability to get at most three different numbers (with repetitions)? $a+b$;

a: "only one number"; c: "only 1's and 2's, but at least a 1 and a 2."

$a=6\left(\frac{1}{6}\right)^{20}$

$b=\binom{6}{2}\cdot c$

$c=\left(\frac{2}{6}\right)^{20}-2\cdot\left(\frac{1}{6}\right)^{20}$

Exercise K. An online cooking videogame randomly assigns 9 players in different roles: 1 player is doing the dishes, 3 are cutting meat and vegetables, 3 are supervising the cooking, and 2 are taking food from place to place. The task partitioning is uniform among partitions satisfying the above requirements.

1. You and your buddy are playing that game. What is the probability that you end up in the same station? $\left[\binom{7}{1,1,3,2}+\binom{7}{1,3,1,2}+\binom{7}{1,3,3}\right]\Big/\binom{9}{1,3,3,2}$

2. Two groups of 4 and 2 people are playing and end up in the same game. What is the probability that for both teams, the players are all isolated? $4!\cdot 2!\cdot 2\Big/\binom{9}{1,3,3,2}$

Independence, conditioning

Textbook: 1.17, 1.18, 1.19, 1.20, 1.21, 1.22, 1.24, 1.25, 1.26, 1.27, 1.29 to 1.39.
Homework III-1.

Exercise L. Two events A and B are defined on some probability space. Which of the following probabilities are compatible with the axioms of probability? (There may be several good answers.)

1. $\mathbb{P}(A) = 50\%$, $\mathbb{P}(B) = 50\%$, $\mathbb{P}(A|B) = 50\%$. Compatible.
2. $\mathbb{P}(A) = 50\%$, $\mathbb{P}(B) = 20\%$, $\mathbb{P}(A|B) = 50\%$. Compatible.
3. $\mathbb{P}(A) = 10\%$, $\mathbb{P}(B) = 20\%$, $\mathbb{P}(A|B) = 80\%$. Incompatible $(\mathbb{P}(A) < \mathbb{P}(A|B)\cdot\mathbb{P}(B))$.
4. $\mathbb{P}(A) = 10\%$, $\mathbb{P}(B) = 50\%$, $\mathbb{P}(A|B) = 50\%$. Incompatible $(\mathbb{P}(A) < \mathbb{P}(A|B)\cdot\mathbb{P}(B))$.

Exercise M. We throw three fair dice.

one 1, one 6, something else.

1. What is the probability $\mathbb{P}(\text{At least two 1} \,|\, \text{The minimum is 1})$? $1-3\cdot\left(\frac{5}{6}\right)^2$
2. What is the probability $\mathbb{P}(\text{At least one 6} \,|\, \text{The minimum is 1})$? $3\cdot\left(\frac{1}{6}\right)^3 + \binom{3}{2}\cdot\frac{1}{6}\cdot\frac{1}{6}\cdot\frac{4}{6}$

two 1's, one 6

Exercise N. We sample a random result X in the following fashion. We throw a fair die once. If it lands of a number from 1 to 5, the result X is the value of a second independent throw. If it lands on 6, we make a second independent throw, and the result X is 7 if the second throw is a number from 1 to 5, otherwise it is $\star$.

What are the probabilities of X being 1, 2, 3, 4, 5, 6, 7, $\star$? $\mathbb{P}(X=1) = \cdots = \mathbb{P}(X=7) = \frac{5}{36}$

$\mathbb{P}(\star = X) = \frac{1}{36}$

Given a biased coin that lands on heads with probability $0 < p < 1$, we sample a random result X in the following fashion. We toss the coin once first. The result X is the value of that first coin provided a second independent throw gives the opposite outcome. Otherwise, it is $\star$.

What are the probabilities of X being heads, tails, $\star$?

$\mathbb{P}(X=H) = \mathbb{P}(X=T) = p(1-p)$

$\mathbb{P}(X=\star) = p^2+(1-p)^2$

(This last question is too opinion-based to be a midterm-type question.)

How would you emulate a fair 7-sided die with a single fair 6-sided die? How would you emulate a fair coin with a single unfair coin? You may throw them many times.

Do the experiment, if you got a star, do it again until you have a valid result.

Law of total probability, Bayes law

Textbook: 1.40, 1.42, 1.43, 1.45, 1.46, 1.47, 1.48, 1.49, 1.51.
Homework: III-1.

Exercise O. 100 black-spotted white cows in a herd have random patterns around the eyes. We assume a black spot cannot cover both eyes. Which of the following collections of events are partitions? (There may be several good answers.)

1. $B_1, \ldots, B_{100}$, where B_k is the event that the kth cow has a black spot around at least one eye. Disjoint, Cover $\Rightarrow$ Partition
2. $B_0, \ldots, B_{100}$, where B_k is the event that at least k cows have a black spot around at least one eye. Disjoint, Cover $\Rightarrow$ Partition
3. $B_0, \ldots, B_{200}$, where B_k is the event that there are exactly k black spots around the eyes of all the cows.

   Disjoint, Cover $\Rightarrow$ Partition

Note the importance of B_0.

4. $B_0, \dots, B_{100}$, where B_k is the event that the kth cow is the first to have a black spot around at least one eye (B_0 is the event that none of them has such a spot). Disjoint, cover $\Rightarrow$ partition
5. $B_0, \dots, B_{100}$, where B_k is the event that all cows up the the kth one have a black sport on at least one eye, and the $(k+1)$st doesn't (B_0 is the event that the first cow does not have such a spot). Disjoint, cover $\Rightarrow$ partition

Exercise P. A swimmer and a runner are competing in a television show involving two events: one on track and one in pool. The swimmer has a 40% chance to pass the track event, and a 60% chance to pass the pool event, independently from each other. The runner has the same probabilities, but the other way around.

We choose a first competitor uniformly at random, and watch them attempt the two trials.

1. Are the event 'The first competitor passes the track event' and 'The first competitor passes the pool event' independent? $\mathbb{P}(\text{passes track \& passes pool}) = 24\%$

   $\mathbb{P}(\text{passes track}) \times \mathbb{P}(\text{passes pool}) = 25\%$ Not independent
2. What is the probability that the first competitor is the swimmer, knowing ~~the~~ that they passed the pool trial? $\mathbb{P}(\text{swimmer} \mid \text{passed pool}) = 60\%$
3. What is the probability that the first competitor is the swimmer, knowing that they passed both trials? $\mathbb{P}(\text{swimmer} \mid \text{passed both}) = 50\%$

Exercise Q. Among the population of students being awarded a bachelor's degree in mathematics in the U.S. in a generation, we found the following statistics.[3] About 8% of them went on to graduate with a PhD, and 5% of these remaining people were granted the title of professor.[4] Out of all the graduates, about 43% were women; among the doctors, we find 29% of women, and it goes down to 15% for the professors.

$\mathbb{P}(\text{professor}) = 0.4\%$

1. What is the probability that a randomly chosen graduate raised to a professor position?
2. What is (a formula for) the probability that a randomly chosen graduate became a professor, knowing that she is a woman? Same question for a man. $\mathbb{P}(\text{professor} \mid \text{woman}) = \frac{15\% \cdot 0.4\%}{43\%}$
3. Complete the following sentence: $\mathbb{P}(\text{professor} \mid \text{man}) = \frac{85\% \cdot 0.4\%}{57\%}$

   'The social situation in academia makes male graduate in maths 4.3 times more likely to be awarded the title of professor than their female counterpart.'

$$\frac{\mathbb{P}(\text{professor} \mid \text{man})}{\mathbb{P}(\text{professor} \mid \text{woman})} = \frac{85 \cdot 43}{57 \cdot 15} \simeq 4.3$$

Discrete probability distributions

Textbook: 2.1, 2.2, 2.3, 2.4.

[3] Based on data from the National Science Foundation, and a few obviously false assumptions; in particular, the fact that the proportion of women is more or less constant.

[4] We are talking about actual professor, the highest rank in the U.S. academia. It is a very prestigious position.

At the end of each add "$\mathbb{P}(N=x)=0$ otherwise"

Exercise R. For each of the following discrete random variables N, find the probability mass function.

ages of the g.

$\mathbb{P}(N=k) = \frac{\binom{3}{k}}{2^3-1}$ for $k=1,2,3$. all boys

1. Select three people at random in Notre Dame. For each day of the week, count how many of them were born on this date. N is the maximum of these 7 numbers. $\mathbb{P}(N=1) = \frac{(7)_3}{7^3}$ $\mathbb{P}(N=3) = \frac{7}{7^3}$

2. An ad company calls random families with 3 children, at least one of which is a girl. N is the number of girls among the children. $\mathbb{P}(N=2)$ $(=1-\mathbb{P}(N=1)-\mathbb{P}(N=3))$

3. A cake is 10% dough, 60% filling, 30% topping. The filling can be an apple sauce (90% fruit), a cream with strawberries (50% fruit), or salted caramel (0% fruit); the topping can be a chocolate ganache (0% fruit) or a fruit salad (100% fruit). The dough is a classic crust (no fruit). A cake is selected uniformly at random among all six possible combinations. N is the proportion of fruit in the cake. $\mathbb{P}(N=0.3)=\frac{1}{3}$ $\mathbb{P}(N=0)=\frac{1}{6}=\mathbb{P}(N=0.6)=\mathbb{P}(N=0.54)=\mathbb{P}(N=0.84)$

L1 L2 L3 L4 L5

S1 S2 S3

4. There are 3 independent secrets in a videogame, that you have a 30% chance of finding and that will grant you four stars each. Upgrades for your gear (initially of level zero) will cost 1 star for the first level, 2 for the second (i.e. a total of 3 stars), 3 stars for the third and so on. N is the level of your gear after you played the game. $\mathbb{P}(N=0)=0.7^3$, $\mathbb{P}(N=2)=3\cdot 0.3\cdot 0.7^2$, $\mathbb{P}(N=1)=0$, $\mathbb{P}(N=3)=3\cdot 0.7\cdot 0.3^2$, $\mathbb{P}(N=4)=0.3^3$

Exercise S. A company sells 1,000 mechanical pieces to a client. Each of these pieces may be defective or not. For each of the following numbers N, give the name and parameter(s) of the corresponding distribution, and (a formula for) the probability $\mathbb{P}(N=5)$.

1. For this question and the next one, assume that exactly two pieces are defective. The client will use only 850 pieces, chosen uniformly at random. N is the number of unfit pieces they will use. HyperGeom$(2, 998, 850)$

2. As above, assume exactly two pieces are defective. The client has a quality control service, but it is incredibly lazy. They will check only one piece at random, and say that the whole batch is either perfect or dreadful, according to the condition of the piece. N is the number of defects that quality control will detect. Ber$(2/1000)$

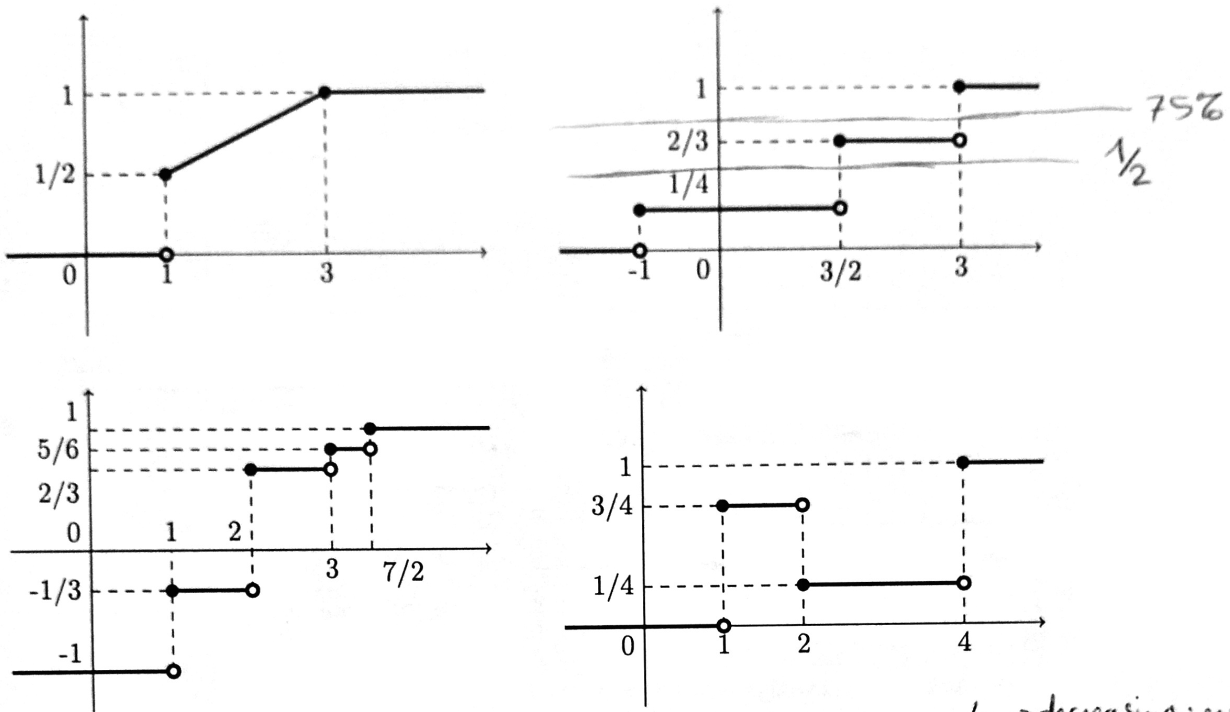
3. For the rest of the exercise, assume that each piece is defective with a probability 0.1%, independently of the others. N is the number of items with a defect. Bin$(1000, 0.001)$

4. For the next contract (same model, same quantity, same probability of defect), the seller wants to be sure that no piece is defective. She will personally inspect each and every item, and throw away the defective ones. N is the number of pieces, defective or not, that the company will produce. NegBin$(1000, 0.999)$

Cumulative distribution function

Textbook: 2.5.

Exercise T.



4 → decreasing : no.

1. Which of the above graphs is the cumulative distribution function of a discrete random variable, and why? 1 → nonzero slope : no.   3 → negative probability : no.
2. Let us call N this variable. What are the probabilities $\mathbb{P}(N = 1)$, $\mathbb{P}(N = 3)$, $\mathbb{P}(N \geq 1)$?

   (In the midterm, there will be no question depending crucially on the previous one.)

   $\mathbb{P}(N=1)=0$, $\mathbb{P}(N=3)=1/3$

   $\mathbb{P}(N\geq 1)=(1/3-1/4)$ $=3/4$

   $+(1-\ /\ )$
3. What is the median of N? What is its 75th percentile?

   (Same remark.) median $= 3/2$

   75th percentile = upper quartile
   $= 3$.