

Exercise 1. Figure 1 shows a graph with 9 edges and 8 vertices, two of which are denoted by A and B . In an experiment, we delete each edge independently with probability $\frac{1}{2}$.

- (a) Write down the sample space Ω for this experiment. What is the cardinality of Ω ?
- (b) With what probability is there a path connecting A and B ?
- (c) With what probability is there a path connecting A and B , if we also know that at most one edge is deleted?

Solution:

(a - 2 points) Let the graph in Figure 1 be denoted by $G = (V, E)$. Then, $\Omega = \{G' = (V, E') \mid E' \subseteq E\}$. The cardinality of Ω is the number of subsets E' of E . Since E has cardinality 9, the number of possible subsets is $2^9 = 512$.

(b - 3 points) Let $A \leftrightarrow B$ denote the event that nodes A and B are connected by a path.

$$\begin{aligned}
 P(A \leftrightarrow B) &= P[e_1 \cap e_9 \cap \{e_5 \cup (e_2 \cap e_4 \cap e_7) \cup (e_3 \cap e_6 \cap e_8)\}] \\
 &= P(e_1)P(e_9) [1 - P\{\bar{e}_5 \cap (\bar{e}_2 \cup \bar{e}_4 \cup \bar{e}_7) \cap (\bar{e}_3 \cup \bar{e}_6 \cup \bar{e}_8)\}] \\
 &= \frac{1}{2} \cdot \frac{1}{2} \cdot [1 - P(\bar{e}_5)P(\bar{e}_2 \cup \bar{e}_4 \cup \bar{e}_7)P(\bar{e}_3 \cup \bar{e}_6 \cup \bar{e}_8)] \\
 &= \frac{1}{4} \cdot [1 - P(e_5) \{1 - P(e_2 \cap e_4 \cap e_7)\} \{1 - P(e_3 \cap e_6 \cap e_8)\}] \\
 &= \frac{1}{4} \cdot \left\{ 1 - \frac{1}{2} \left(1 - \frac{1}{2^3} \right)^2 \right\} \\
 &= \frac{79}{512}.
 \end{aligned}$$

(c - 3 points) Let m be the number of edges in the new graph G' . Then, we need to compute

$$P(A \leftrightarrow B \mid m \geq 8) = \frac{P(A \leftrightarrow B, m \geq 8)}{P(m \geq 8)}.$$

There are ten possible G' with $m \geq 8$; one for the case when all the edges are present and nine for the cases when one of them are absent. Since all possible G' has equal $\frac{1}{|\Omega|}$ probability,

$$P(m \geq 8) = \frac{10}{512}.$$

Out of these, A and B are connected in all but two cases; the case when e_1 is missing and when e_9 is missing. Again, all these G' have $\frac{1}{|\Omega|}$ probability, so

$$P(A \leftrightarrow B, m \geq 8) = \frac{10 - 2}{512} = \frac{8}{512}.$$

Putting it all together,

$$P(A \leftrightarrow B \mid m \geq 8) = \frac{\frac{8}{512}}{\frac{10}{512}} = \frac{4}{5}.$$

Exercise 2. According to the statistics department, about 1% of the students cheat during exams at EPFL. To detect cheating students more accurately, EPFL designed a new method based on an AI system. The detector can correctly detect a cheater with 90% probability, but unfortunately, the detector classifies a honest student as a cheater with 2% probability.

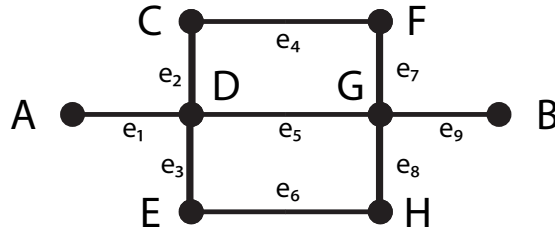


Figure 1: The graph for Exercise 1.

- (a) The detector said your friend Alice was cheating on the exam. What is the probability that she actually cheated? Formulate the problem described above precisely by defining random variables and conditional probabilities.
- (b) The experiment can be repeated n times for the same student, and the detector's answer is conditionally independent on the state of the student (cheater or honest). A student is declared a cheater if all of the n times the detector said he or she is cheating. How large does n need to be so that a student declared as a cheater has more than 99.99% chance of being actually a cheater?

Solution:

- (a - 3 points) We denote by "C" the event that a person was cheating on the exam, and "D" the event that the detector said that the person was cheating. We know from the problem statement that

$$P(D|C) = 0.9, \quad P(D|\bar{C}) = 0.02, \quad P(C) = 0.01.$$

This allows us to compute the probability that your friend was cheating if the detector said cheater:

$$P(C|D) = \frac{P(C \cap D)}{P(D)} = \frac{P(D|C)P(C)}{P(D|C)P(C) + P(D|\bar{C})P(\bar{C})} = \frac{0.9 \times 0.01}{0.9 \times 0.01 + 0.02 \times (1 - 0.01)} = \frac{5}{16} = 0.3125.$$

- (b - 3 points) We denote by D_n the event that the detector said cheater n times. Since the answers of the detector are conditionally independent on the state of the student (cheater or honest), $P(D_n|C) = 0.9^n$ et $P(D_n|\bar{C}) = 0.02^n$. So we have

$$P(C|D_n) = \frac{P(D_n|C)P(C)}{P(D_n|C)P(C) + P(D_n|\bar{C})P(\bar{C})} = \frac{0.9^n \times 0.01}{0.9^n \times 0.01 + 0.02^n \times (1 - 0.01)}.$$

This formula gives us $P(C|D_1) = 0.3125$, $P(C|D_2) = 0.95339$, $P(C|D_3) = 0.998915$, $P(C|D_4) = 0.999976 > 0.9999$ So it is enough to repeat the experiment four times.

Remark: It is also possible to find n explicitly. We are looking for the smallest n for which:

$$\begin{aligned} 0.9^n \times 0.01 &\geq 0.9999(0.9^n \times 0.01 + 0.02^n \times 0.99) \\ \Leftrightarrow 0.9^n &\geq 0.9999 \times 0.9^n + 0.9999 \times 0.02^n \times 99 \\ \Leftrightarrow 0.0001 \times 0.9^n &\geq 0.9999 \times 99 \times 0.02^n \\ \Leftrightarrow \left(\frac{0.9}{0.02}\right)^n &= 45^n \geq 9999 \times 99 \\ \Leftrightarrow n \log 45 &\geq \log(9999 \times 99) \\ \Leftrightarrow n &\geq \frac{\log(9999 \times 99)}{\log 45} \approx 3.626631. \end{aligned}$$

Hence we find that it is enough to repeat the experiment 4 times.

Exercise 3. Consider rolling two fair dice, and consider the addition of their scores as the random variable X . Assume you have two options for gambling: 1. If $X = 7$, you will win 1 CHF, otherwise you will loose B_1 CHF; 2. If $X = 2$ or $X = 12$, you will win 1 CHF, otherwise you will loose B_2 CHF.

- (a) For each option, compute the probability of winning. We call them P_1 and P_2 .
- (b) For each option, compute the expected value of the money you will win as functions of B_1 and B_2 . We call them μ_1 and μ_2 .
- (c) For each option, compute the variance of the money you will win as functions of B_1 and B_2 . We call them σ_1^2 and σ_2^2 .
- (d) Find B_2 as a function of B_1 such that $\mu_1 = \mu_2$. We call it B_2^* .
- (e) *BONUS:* When $B_2 = B_2^*$, find σ_2^2 as a function of σ_1^2 . Which one is smaller? In this case, if you want to maximize the probability of winning, which option is better for you? If you want to minimize the variance, which option is better for you?

Solution:

(a - 2 points)

$$P_1 = P[X = 7] = 6 \frac{1}{36} = \frac{1}{6}$$

$$P_2 = P[X = 2] + P[X = 12] = 2 \frac{1}{36} = \frac{1}{18}$$

(b - 3 points)

$$\mu_1 = \frac{1}{6}1 - \left(1 - \frac{1}{6}\right)B_1 = \frac{1 - 5B_1}{6}$$

$$\mu_2 = \frac{1}{18}1 - \left(1 - \frac{1}{18}\right)B_2 = \frac{1 - 17B_2}{18}$$

(c - 3 points) One way for this part is to basically write the equation and try to simplify it. The other one is to consider the outcome of each game as a Bernoulli random variable, and then we will have:

$$\sigma_1^2 = \frac{1}{6}\left(1 - \frac{1}{6}\right)(1 + B_1)^2 = \frac{5}{36}(1 + B_1)^2$$

$$\sigma_2^2 = \frac{1}{18}\left(1 - \frac{1}{18}\right)(1 + B_2)^2 = \frac{17}{18^2}(1 + B_2)^2$$

(d - 2 points)

$$B_2^* = \frac{15B_1 - 2}{17}$$

(e - 1 point) By substituting B_2^* in the results of part c:

$$\sigma_2^2 = \frac{17}{18^2}\left(1 + \frac{15B_1 - 2}{17}\right)^2 = \frac{17}{18^2} \frac{15^2}{17^2}(1 + B_1)^2 = \frac{5}{17}\sigma_1^2$$

As it can be seen, no matter what happens, σ_1^2 and P_1 are respectively larger than σ_2^2 and P_2 . So, if one wants to maximize the winning probability, he or she should choose the 1st option, and if one wants to minimize variance (i.e. risk), he or she should choose the 2nd option.

Note: In the case that one asks, it can be shown that it is feasible to have possible values for B_1 , $B_2 = B_2^*$, μ_1 , and μ_2 at the same time.

Exercise 4.

- (a) A component of a system breaks down independently with probability $p = 0.001$ each day. Compute the probability that the component is still functional after 1000 days.
- (b) Compute $P[X > E[X]]$, where $X \sim \text{Exp}(\lambda)$ is an exponential random variable.
- (c) *BONUS*: Compare the results of part a) and b) and explain it intuitively using a property seen in class.

Solution:

(a - 3 points) Let the day on which the component breaks denoted by Y . The random variable Y follows a geometric distribution.

$$\begin{aligned} P(Y > 1000) &= 1 - P(Y \leq 1000) = 1 - \sum_{i=1}^{1000} p(1-p)^{i-1} = 1 - p \frac{1 - (1-p)^{1000}}{1 - (1-p)} = (1-p)^{1000} = \\ &= \left(1 - \frac{1}{1000}\right)^{1000} \approx e^{-1}, \end{aligned}$$

since $\lim_{n \rightarrow \infty} (1 - x/n)^n = e^{-x}$.

(b - 3 points) Let X be $\text{Exp}(\lambda)$.

$$E[X] = \lambda \int_0^{\infty} x e^{-\lambda x} dx = \lambda \left[-\frac{e^{-\lambda x}(\lambda x + 1)}{\lambda^2} \right]_0^{\infty} = \frac{1}{\lambda}.$$

Now we can directly address the problem,

$$P(X > E[X]) = P(X > 1/\lambda) = \int_{1/\lambda}^{\infty} \lambda e^{-\lambda x} dx = \left[-e^{-\lambda x} \right]_{1/\lambda}^{\infty} = e^{-1}.$$

(c - 1 point) Since $E[Y] = 1000$, both problems ask for the probability of a random variable exceeding its expected value. Moreover the two distributions X and Y are very similar (after some scaling).

We will make this “very similar” more precise, similarly as we did in class for Poisson approximation. There we showed

$$\text{Bin}(n, \lambda/n) \xrightarrow{D} \text{Poi}(\lambda)$$

This time we show

$$\frac{1}{n} \text{Geom}(\lambda/n) \xrightarrow{D} \text{Exp}(\lambda)$$

We can see this by computing the CDF of both distributions. Recall,

$$F_{\text{Exp}(\lambda)}(x) = 1 - e^{-\lambda x}$$

Also, for $k \in \{\frac{1}{n}, \frac{2}{n}, \dots\}$,

$$F_{\frac{1}{n} \text{Geom}(\lambda/n)}(k) = F_{\text{Geom}(\lambda/n)}(nk) = 1 - \left(1 - \frac{\lambda}{n}\right)^{nk} \rightarrow 1 - e^{-\lambda k}$$

Hence $F_{\frac{1}{n}\text{Geom}(\lambda/n)}(x) \rightarrow F_{\text{Exp}(\lambda)}(x)$ is true (for rational x , but we will not discuss the irrational case here), which is the definition for convergence in distribution in the slides (Definition 103).

This derivation is quite challenging and was not required to get full points for this bonus question. 0.5 points were awarded for noticing that the two results are numerically close and 0.5 points were awarded for mentioning convergence or the Law of small numbers or Poisson approximation in the right context.