Indian Institute of Technology Bombay
Department of Electrical Engineering

Handout 8 EE 325 Probability and Random Processes
Tutorial 2 Feb 1, 2019

Question 1) We discussed the construction of a uniform probabiliy on $\Omega = [0,1]^2$. Using those techniques, construct a uniform probability on the unit circle in $\mathbb{R}^2$.
Hint: The idea of restriction of the space by conditioning can be handy.

Question 2) Find the probability that two randomly picked points on a line segment partition it into 3 parts that form a triangle? Before you start computing, first define an appropriate probability space $(\Omega, \mathcal{F}, P)$. Now construct a measurable map suitable for the event of interest. Then compute the probability as an induced measure.
Solution: There are many ways to calculate this probability, we will describe the most general approach. In particular we will take $\Omega = [0,1]^2$, $\mathcal{F} = \mathcal{B}(\mathbb{R})$ and the uniform probability measure on Ω. We have already defined these things in the lectures.

Let X be a random vector with the respective coordinates $(X.x)$ and $(X.y)$ describing two points at which a unit line segment is cut. Which all points in Ω will be such that the three segments formed by the cuts $(X.x)$ and $(X.y)$ can be arranged to a triangle. Well, this is clearly marked in the figure below on Ω.

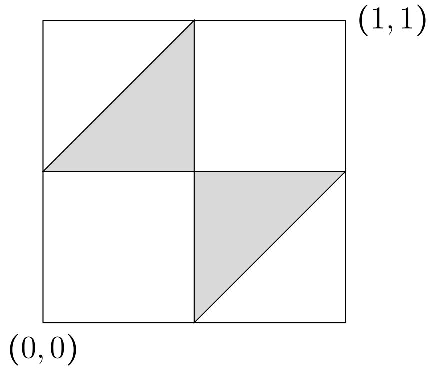

Figure 1: Points on $\Omega = [0,1]^2$ which leads to a triangle

Question 3) For events $A_i, i \geq 1$, show that

$$P(\bigcup_{i=1}^{n} A_i) \geq \sum_{j=1}^{n} P(A_j) - \sum_{1 \leq i < j \leq n} P(A_i A_j).$$

Question 4) Show that any function $X : (\Omega_1, \{\Omega_1, \varnothing\}) \to (\Omega_2, \{\Omega_2, \varnothing\})$ is measurable. How will this result change if the second sigma-field is replaced by $\{\Omega_2, A, A^c, \varnothing\}$ for some meaningful subset A.

Solution: We proceed by induction.

$$\begin{aligned}
P(\bigcup_{i=1}^{n} A_i) &= P(\bigcup_{i=1}^{n-1} A_i) + P(A_n) - P(\bigcup_{i=1}^{n-1} A_i \bigcap A_n) \\
&\geq \sum_{j=1}^{n-1} P(A_j) - \sum_{1\leq i<j\leq n-1} P(A_i A_j) + P(A_n) - P(\bigcup_{i=1}^{n-1} A_i \bigcap A_n) \\
&= \sum_{j=1}^{n} P(A_j) - \sum_{1\leq i<j\leq n-1} P(A_i A_j) - P(\bigcup_{i=1}^{n-1} (A_i \bigcap A_n))
\end{aligned}$$

The second step above used the induction argument for $n-1$. Now applying the union bound

$$\begin{aligned}
P(\bigcup_{i=1}^{n} A_i) &\geq \sum_{j=1}^{n} P(A_j) - \sum_{1\leq i<j\leq n} P(A_i A_j) \sum_{i=1}^{n-1} P(A_i, A_n) \\
&= \sum_{j=1}^{n} P(A_j) - \sum_{1\leq i<j\leq n} P(A_i A_j).
\end{aligned}$$

Question 5) For a discrete random variable X taking values in a countable state-space E, we will call $X^{-1}(x) := \{\omega \in \Omega : X(w) = x\}$ as the events associated with the random variable X.

Definition 1 *Two discrete random variables X_1 and X_2 on E are said to be independent if*

$$P(X_1^{-1}(u), X_2^{-1}(v)) = P(X_1^{-1}(u))P(X_2^{-1}(v)) \forall u, v \in E.$$

i.e. any pair of their respective associated events are independent.

Define independence for two random variables X_1 and X_2 taking values in E_1 and E_2 respectively.
Solution:

$$P(X_1^{-1}(u), X_2^{-1}(v)) = P(X_1^{-1}(u))P(X_2^{-1}(v)) \forall u \in E_1, v \in E_2.$$

Question 6) Consider n random variables $X_1, X_2, \cdots X_n$. We call them independent if the associated events are so (Recall the definition of independence of many events in class).
(a) Show that iff, $\forall u_i \in E_i$,

$$P(\bigcap_{i=1}^{n} X_i^{-1}(u_i)) = \prod_{i=1}^{n} P(X_i^{-1}(u_i)),$$

then $X_1, \cdots, X_n$ are independent random variables. Notice that we only considered products with n terms, unlike our earlier definition for all subsets of $\{1, \cdots, n\}$.
Solution: Done in class.
b) Consider a discrete random variable taking values in $\{1, 2, \cdots, k\}$ with $P(X = i) = p_i, 1 \leq i \leq k$. For n draws of this random variable, say $\bar{X} = X_1, \cdots, X_n$, and for any given sequence $\bar{u} = (u_1, \cdots, u_n)$ it is specified that

$$P(\bar{X} = \bar{u}) = \prod_{i=1}^{k} p_i^{N_i(\bar{u})},$$

where $N_i(\bar{u})$ counts the number of times i appears in the sequence $\bar{u}$. Are the draws in $\bar{X}$ independent.
(c) For the previous part, let us now compute the average,

$$\frac{1}{n}\sum_{l=1}^{n} X_l.$$

Show that

$$\frac{1}{n}\sum_{l=1}^{n} X_l = \sum_{j=1}^{k} a_j \frac{\sum_{l=1}^{n} \mathbb{1}_{\{X_l=j\}}}{n},$$

for some constants $a_j, 1 \le j \le k$. Also find the explicit values of the constants $a_j, 1 \le j \le k$. If you have doubt, just see what happens when $n = 1$, $n = 2$ etc.

Question 7) Consider a random variable X with probability distribution

$$P(X = n) = \frac{c}{n^2 \log n}, \forall n \in \mathbb{N}, n \ge 2$$

where c is an appropriate constant. Compute $\mathbb{E}[X]$. Recall, expectation is defined as the quantity

$$\mathbb{E}[X] = \sum_{x \in E} x P(X = x),$$

whenever the sum is well-defined. How will your answer change if the distribution is taken as $P(X = n) = \frac{d}{n^2(\log n)^2}$ for an appropriate constant d.
Solution: First for all notice that the given distribution is a valid one since the $\sum_{n\ge 2} \frac{1}{n^2 \log n}$ converges a constant, we have denoted this constant by $\frac{1}{c}$. The expectation is

$$\mathbb{E}X = \sum_{n\ge 2} n \frac{c}{n^2 \log n} = \sum_{n \ge 2} \frac{c}{n \log n} = \infty.$$

Thus, this random variable which takes only finite values, i.e. $\mathcal{N}$ has an infinite expectation. When the random variable takes only non-negative values, there is no harm in the expectation being unbounded. We can clearly identify this unbounded number with positive infinity. It is for this reason that the lecture notes defines expectation for all non-negative valued random variables. For general RVs in $\mathbb{R}$ (i.e. both positive and negative values are possible), the integrability condition of $\mathbb{E}g(X) < \infty$ is required for the definition.

As fot the second part, we can use an integral approximation to show that

$$\frac{1}{\log 2} \le \sum_{n \ge 2} n \frac{1}{n^2 (\log n)^2} \le \frac{1}{\log 2} + \frac{1}{\log 3}$$

Given below are some questions for your practice, they are simple enough, and require no particular background

Question 8) There are three boxes: (i) a box containing two gold coins, (ii)a box containing two silver coins, (iii) a box containing one silver coin and one gold coin.

Let us now choose a box randomly and then choose a coin from that box randomly. Suppose it turns out to be a gold coin. What is the probability that the chosen box is the one that contains two gold coins?

Solution: Let us name the respective Boxes as 1, 2, and 3. The gold and Box 1 happens with probability 13, whereas gold and Box II happens with probability $\frac{1}{2}\frac{1}{3}$.

$$P(Box\ 1|Gold) = \frac{\frac{1}{3}}{\frac{1}{3} + \frac{1}{3}\frac{1}{2}} = \frac{2}{3}.$$

Question 9) Consider a MCQ (multiple choice question) based exam. Assume that for each question a student knows (independently) the answer with probability p and guesses it with $1 - p$. Assume that the student's guess is correct with probability $\frac{1}{m}$, where m is the number of MCQs. What is the conditional probability that a student knew the answer to a question given that she answered it correctly.
Solution: The correct answer can come from either knowing or guess-work, the latter having probability of success $\frac{1-p}{m}$. The probability in question now is

$$P(Know|Correct) = \frac{p}{p + \frac{1-p}{m}}$$

Question 10) A pond contains 3000 red and 7000 golden fishes, of which 200 and 500, respectively, are tagged. Find the probability that a random sample of 100 red and 200 golden fishes will show 15 and 20 tagged red and golden fishes respectively.
Solution:

$$\frac{\binom{200}{15}\binom{2800}{85}\binom{500}{20}\binom{7500}{180}}{\binom{3000}{100}\binom{7000}{200}}$$