# Online Resit (2020-2021) Probability Theory for EOR

–

# Question 1 –3

## See additional exercises discussions at the end

(10 points) This question consists of two unrelated subquestions.

(a) Four kids are ringing doorbells in your neighborhood. There are 12 doorbells corresponding to the numbers $1, 2, \ldots, 12$. Each kid randomly presses one doorbell. For example, they may ring doorbells $\{2, 1, 3, 4\}$, $\{1, 1, 8, 9\}$ or $\{6, 6, 6, 12\}$.

What is the probability that the doorbells corresponding to four successive even numbers, for example $\{2, 4, 6, 8\}$, are rung?

Chapter 1 - Naive Definition of Probability.

Denote by $S$ the event that the kids ring doorbells corresponding to four successive even numbers. Denote by $N_T$ the total number of ways in which the kids can select the doorbells. The first kid has 12 choices, the second kid has again 12 choices, and the same for the third and fourth kid. By the multiplication rule, there are $N_T = 12^4$ ways the kids can select the doorbells.

Now we turn to the favorable outcomes. There are three sequences of successive even numbers, namely $\{2, 4, 6, 8\}$, $\{4, 6, 8, 10\}$ and $\{6, 8, 10, 12\}$. Each of these can be ordered in 4! ways, so using the MR, we have $N_D = 3 \cdot 4!$ ways in which the kids can ring 4 successive even doorbells.

Using the naive definition of probability, we get

$$P(S) = \frac{N_D}{N_T} = \frac{3 \cdot 4!}{12^4} = 0.0035.$$

**(a) - Version 2** Consider drawing one by one from a well-shuffled standard deck of cards. A standard deck of cards contains 52 cards. A card has a rank (which can be 2-10, Jack, Queen, King, or Ace), and a suit (diamonds, hearts, spades or clubs).

We draw a *street* if we draw four cards that can be arranged in consecutive ranks, for example we draw $\{2, 3, 4, 5\}$, $\{2, 3, 5, 4\}$, or $\{J, Q, K, 10\}$.

Suppose that the fifth card you draw is your first Ace. What is the probability of seeing a street before this first Ace?

Chapter 1 - Naive Definition of Probability.

Denote by $S$ the event that we see a street before the first ace on the fifth card.

First, consider the total number of ways to select the four cards before the first ace on the fifth. Given the fact that we did not draw any Ace's in the first four cards, we know the first four cards are drawn from 48 cards. In total, there are $N_T = \binom{48}{4}$ ways to draw the four cards prior to the first ace.

Now we turn to the favorable outcomes. We can have $\{2, 3, 4, 5\}$ up to $\{10, J, Q, K\}$, so there are 9 different streets. Each street can be chosen in $4^4$ ways (4 ways to select the 2, 4 ways to select the 3, etc). Using the MR, we have $N_S = 4^4 \cdot 9$ ways to draw a street before the first ace.

Using the naive definition of probability, we then have

$$P(S) = \frac{N_S}{N_T} = \frac{4^4 \cdot 9}{\binom{48}{4}} = 0.0118.$$

(b) Consider a random variable $R$ such that $P(R = 1) = 1/2$ and $P(R = -1) = 1/2$. We also have a random variable $X$, which is independent of $R$. Suppose $P(RX < 0) = P(X < 0)$. Show that $P(X < 0) = P(X > 0)$.

Chapter 2: LOTP and independence.

We write out $P(RX < 0)$ using the LOTP where we condition on the value of $R$,

$$\begin{aligned} P(RX < 0) &= \sum_{r=\{-1,1\}} P(RX < 0|R = r)P(R = r) && \text{(LOTP)} \\ &= \frac{1}{2}(P(X < 0) + P(-X < 0)) && \text{(Independence of } R \text{ and } X\text{)} \\ &= P(X < 0) && \text{(Given in the question)} \end{aligned}$$

From the last two lines, we obtain

$$2P(X < 0) = P(X < 0) + P(-X < 0).$$

Hence,

$$P(X < 0) = P(-X < 0).$$

Also, $P(-X < 0) = P(X > -0) = P(X > 0)$, so we now have

$$P(X < 0) = P(X > 0).$$

**(b) - Version 2** Consider a random variable $R$ such that $P(R = 1) = 1/2$ and $P(R = -1) = 1/2$. We also have a random variable $X$, which is independent of $R$. Suppose $P(X < 0) = P(X > 0)$. Show that $P(RX < 0) = P(X < 0)$.

Chapter 2: LOTP and independence.

We write out $P(RX < 0)$ using the LOTP where we condition on the value of $R$,

$$\begin{aligned} P(RX < 0) &= \sum_{r=\{-1,1\}} P(RX < 0|R = r)P(R = r) && \text{(LOTP)} \\ &= \frac{1}{2}(P(X < 0) + P(-X < 0)) && \text{(Independence of } R \text{ and } X\text{)} \\ &= P(X < 0). \end{aligned}$$

were the last line uses that $P(-X < 0) = P(X > -0) = P(X > 0)$ and the fact that it is given that $P(X < 0) = P(X > 0)$.

(15 points) **Please note: different versions contained different phrasings of this problem, but the solutions are identical.**

**Consider a network of $n$ people that includes you and your friend. On a given day, $k$ randomly selected pairs have a phone call. The pairs are drawn**

**without replacement. As an example, when $k = 2$ we could have the pairs $\{$You, Your friend$\}$ and $\{$You, someone other than You or Your friend$\}$. We make no distinction between outgoing and incoming calls.**

(a) What is the probability that you have a phone call with your friend?

Chapter 1: Naive definition of probability.

Denote by $p$ the probability that you talk to you friend. The total number of possible pairs is $N = \binom{n}{2}$. It is given in the question that we draw from these pairs without replacement.
The number of selected pairs where you talk to your friend is $N_F = \binom{1}{1}\binom{N-1}{k-1}$. The total number of ways to select the pair is $N_T = \binom{N}{k}$.

Using the naive definition of probability, we conclude that

$$\begin{aligned} p &= \frac{N_F}{N_T} \\ &= \frac{\binom{N-1}{k-1}}{\binom{N}{k}} = \frac{\frac{(N-1)!}{(k-1)!(N-k)!}}{\frac{N!}{k!(N-k)!}} = \frac{(N-1)!}{N!}\frac{k!}{(k-1)!} = \frac{k}{N} = \frac{2k}{n(n-1)}, \qquad 0 \leq k \leq N. \end{aligned}$$

**Alternative solution** You can also use indicator functions here. Denote by $X_i$ the random variable equal to 1 if the $i$-th draw is you and your friend and 0 otherwise. Then, $\mathrm{E}[X_i] = \frac{1}{N}$. Consider $Y = \sum_{i=1}^{k} X_i$. You talk to your friend if $Y = 1$ and you don't talk to your friend if $Y = 0$. Now,

$$\begin{aligned} P(Y = 1) = \mathbb{E}[Y] &= \mathbb{E}[\sum_{i=1}^{k} X_i] \\ &= \sum_{i=1}^{k} \mathbb{E}[X_i] && \text{(Linearity)} \\ &= \frac{k}{N} = \frac{2k}{n(n-1)}, \qquad 0 \leq k \leq N. \end{aligned}$$

(b) What is the probability that a randomly selected person (not being you or your friend) has a phone call with you and a phone call with your friend?

Denote the event that person $i$ calls both you and your friend by $A_i$.

Divide the total number of pairs into a set $\{i, You\}$, $\{i, Your friend\}$ and the set containing all remaining pairs. For a favorable outcome, we need to select 1 out of 1 item in the first set, 1 out of 1 item in the second set, and $k-2$ out of $N-2$ items in the third set. So the total number of favoralbe outcomes is $N_F = 1 \cdot 1 \cdot \binom{N-2}{k-2}$.

The total number of ways to select the pairs is again $N_T = \binom{N}{k}$ So the probability that a randomly selected person talks to both you and your friend is

$$P(A_i) = \frac{N_F}{N_T} = \frac{\binom{N-2}{k-2}}{\binom{N}{k}} = \frac{k(k-1)}{\binom{n}{2}(\binom{n}{2}-1)}.$$

**Alternatively**: We know from (a) that a specific pair calls with probability $k/\binom{n}{2}$. The probability of having two specific pairs where the second is different from the first, is than $P(B_1 \cap B_2) = P(B_1)P(B_2|B_1) = \frac{k}{N}\frac{k-1}{N-1}$, where $B_j$ is the event that we select specific pair $j$ (for example $j = 1$ is me and person $i$, and $j = 2$ is my friend an person $i$).

(c) Let $X$ be the random variable describing the number of people you have a phone call with. Derive the cumulative distribution function (CDF) of $X$ for general $n$ and all possible values of $k$. Draw the CDF for $n = 4$ and $k = 3$.

Chapter 3: Hypergeometric r.v. and CDF.

We first note that for any $n$, the possible values of $k$ are $\{0, \ldots, \binom{N}{2}\}$.

We first derive the probability mass function (PMF). Notice that you don't call anybody if all $k$ pairs are selected from the $N-(n-1)$ pairs that do not include you, so $P(X=0) = \frac{\binom{N-(n-1)}{k}}{\binom{N}{k}}$.

Similarly, if you call one person, this means that we select 1 element from the possible $n-1$ pairs that include you and $k-1$ elements from the remaning $N-(n-1)$ pairs.

Continuing with this argument, you call $r$ persons with probability

$$P(X=r) = \frac{\binom{n-1}{r}\binom{N-(n-1)}{k-r}}{\binom{N}{k}}.$$

This is the hypergeometric distribution with $n-1$ white balls (corresponding to phone calls with you) and $N-(n-1)$ black balls (corresponding to phone calls not with you) and you draw $k$ balls, so $X \sim \text{Hgeom}(n-1, N-(n-1), k)$.

For the CDF, notice that the maximum number of phone calls that you can get is $\min(k, n-1)$. We then have for $0 \le r \le \min(k, n-1)$,

$$P(X \le r) = \sum_{m=0}^{r} P(X=m) = \sum_{m=0}^{r} \frac{\binom{N-(n-1)}{k-m}\binom{n-1}{m}}{\binom{N}{k}}.$$

For $n=4$ and $k=3$, we find $F_X(0) = 1/20$, $F_X(1) = 1/2$, $F_X(2) = 19/20$ and $F_X(3) = 1$.

For $n=5$ and $k=3$, we find $F_X(0) = 1/6$, $F_X(1) = 2/3$, $F_X(2) = 29/30$ and $F_X(3) = 1$.

The figure should clearly show the three properties of a valid CDF: increasing, appropriate limits, right-continuity.

(20 points) **Please note: different versions use slightly different notation, but the solutions are identical.**

A group of $n \ge 4$ persons independently draw numbers from $\text{Unif}(0,1)$. The person who has the highest number wins our textbook free of charge. Let $X_i$ be the number drawn by the $i$th individual, so that $X_i \sim_{i.i.d.} \mathrm{Unif}(0,1)$.

(a) What is the probability of having strictly more than one winner?

Chapter 5: continuous random variables.

(1) The probability of having strictly more than one winner is zero (**use words to motivate is also okay, here I only show solutions motivated by formulas** Methods are not unique, using indicator functions is also okay.):

$$\mathbb{P}\left(\left\{\bigcup_{i=1}^{n}\left\{\max_{l\neq i}\{X_l\} = \max_{1\leq j\leq n}\{X_j\}\right\}\right\}^c\right) \leq \mathbb{P}\left(\bigcup_{i\neq j}\{X_i = X_j\}\right) \leq \sum_{i\neq j}\mathbb{P}\left(X_i = X_j\right) = 0$$

where the last equation is due to the fact that $X_i - X_j$ is one continuous r.v. and thus $\mathbb{P}\left(\{X_i = X_j\}\right) = \mathbb{P}\left(\{X_i - X_j = 0\}\right) = 0$ (the probability of one continuous r.v. equal to one fixed constant is zero).

(b) What is the probability of the $j$th person having the largest number among the first $j$ individuals?

Chapter 5: continuous random variables

By symmetry of continuous random variables we know

$$\mathbb{P}\left(X_{a_1} < X_{a_2} < \cdots < X_{a_j}\right) = 1/j!$$

for arbitrary permutation $(a_1, \cdots, a_j)$ of $(1, \cdots, j)$. Now among all permutations, there are $(j-1)!$ permutations of the format $(a_1, \cdots, a_{j-1}, j)$: therefore,

$$\mathbb{P}\left(\bigcup_{(a_1,\cdots,a_{j-1})}\{X_{a_1} < X_{a_2} < \cdots < X_{a_{j-1}} < X_j\}\right) = 1/j$$

(c) Calculate the expectations: $\mathbb{E}\left(\log\left(\frac{X_1}{1-X_1}\right)\right)$ and $\mathbb{E}\left(\log\left(\frac{X_1+X_2}{2-(X_{n-1}+X_n)}\right)\right)$.

Chapter 4: Expectation + Chapter 5: Continuous random variables.

Note that $X_1$ follows the same distribution as $1-X_1$ (they have the same probability density function (PDF)), and $X_1+X_2$ follows the same distribution as $1-X_{n-1}+1-X_n$ (they have the same PDF). Therefore, by the LOTUS,

$$\mathbb{E}\left(\log\left(\frac{X_1}{1-X_1}\right)\right) = \mathbb{E}\left(\log\left(X_1\right)\right) - \mathbb{E}\left(\log\left(1-X_1\right)\right)$$
$$= \int_{-\infty}^{+\infty} \log(x)f(x)dx - \int_{-\infty}^{+\infty} \log(x)f(x)dx = 0$$

and

$$\mathbb{E}\left(\log\left(\frac{X_1+X_2}{2-(X_{n-1}+X_n)}\right)\right) = \mathbb{E}\left(\log\left(X_1+X_2\right)\right) - \mathbb{E}\left(\log\left(2-(X_{n-1}+X_n)\right)\right)$$
$$= \int_{-\infty}^{+\infty} \log(x)g(x)dx - \int_{-\infty}^{+\infty} \log(x)g(x)dx = 0$$

(d) Derive the moment generating function (MGF) of $Y = a + bX_1$.

Chapter 6: MGF.

We first derive the MGF of $X_1$ (for $t \neq 0$):

$$\mathbb{E}[e^{tX_1}] = \int_0^1 e^{tx} dx = (e^t - 1)/t$$

for $t = 0$, we know $\mathbb{E}[e^{tX_1}] = \mathbb{E}[1] = 1$. Therefore the MGF is well defined (**by the fact that** $\lim_{t \to 0} M_X(t) = 1$). Next, for $t \neq 0$

$$M_{a+bX_1}(t) = e^{at} M_{X_1}(bt) = e^{at} \left(e^{bt} - 1\right) / (bt)$$

for $t = 0$, $M_{a+bX_1}(t) = 1 (= \lim_{t \to 0} e^{at} \left(e^{bt} - 1\right) / (bt))$.

(As long as you derive both t=0 and t=1 cases, you will have full marks, so the discussion part in blue is not part of the grading scheme, but it is one essential part to complete the answer.)

To complete the derivation of the MGF, we always need to show they are well defined, which is to show that there exits a interval $(-a, a)$ such that for $t \in (-a, a), M(t) < \infty$.

This is true because $\lim_{t \to 0} e^{at} \left(e^{bt} - 1\right) / (bt) = 1 = M(0)$.

$\lim_{t \to 0} e^{at} \left(e^{bt} - 1\right) / (bt) = 1 = M(0)$ means that for a fixed arbitrary $\epsilon > 0$, we can find $a_\epsilon > 0$ such that for $|t| < a_\epsilon$, $|M(t) - 1| < \epsilon$ and thus we know for $t \in (-a_\epsilon, a_\epsilon), M(t) < 1 + \epsilon < \infty$.

# Week 7 additional exercises

Here I only list some common and well known results (items in color gray are some extra discussions) that are covered in this course.

- **Properties of CDF.**

  - Increasing from zero to one:
    - Increasing: if $x_1 \leq x_2$, $F(x_1) \leq F(x_2)$.
    - Convergence to zero and one in the limits: $\lim\limits_{x \to -\infty} F(x) = 0$, $\lim\limits_{x \to \infty} F(x) = 1$.

    The above two properties also imply that $0 \leq F(x) \leq 1$ for all $x$.
  - Right-continuous: for any $x$, $F(x) = \lim\limits_{a \to x^+} = F(x)$ (or sometimes, we also use equivalent notation: $\lim\limits_{a \downarrow x} F(a) = F(x)$ where $a$ approaches $x$ from right.)

    The property means sometimes, we may observe $F(x) = \lim\limits_{a \to x^-} \neq F(x)$ (or $\lim\limits_{a \uparrow x} F(a) \neq F(x)$) for some $x$.

- **Properties of PMF.**

  - Non-negativity: $p_X(x) > 0$ for $x \in \mathit{support}(X)$, and $p_X(x) = 0$ otherwise;
  - sum to one: $\sum_{x \in \mathit{support}(X)} p_X(x) = 1$.

    The above two properties also imply that $p_X(x) \leq 1$ for all $x$.

- **Properties of PDF.**

  - Non-negativity: $f_X(x) \geq 0$ s.t. $f_X(x) > 0$ for $x \in \mathit{support}(X)$, and $f_X(x) = 0$ otherwise;
  - integrates to one: $\int_{-\infty}^{\infty} f_X(x)dx = \int_{x \in \mathit{support}(X)} f_X(x)dx = 1$.

    Note here we can have cases where $f_X(x) > 1$.

- **Write down the PDFs and CDFs of Bern, Bin, Expo, Pois, Normal.**

- Bern(p):
  * PMF: $p_X(0) = 1 - p.p_X(1) = p$ and 0 otherwise. (or $p_X(x) = pI_{\{1\}}(x) + (1 - p)I_{\{0\}}(x)$).
  * CDF: $F_X(x) = \sum_{a<x} p_X(a) = pI_{[1,\infty)}(x) + (1-p)I_{[0,\infty)}(x)$.
- Bin(n,p):
  * PMF: $p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}, k = 0, 1, 2, 3, \cdots, n$ and 0 otherwise.
  * CDF: $F_X(x) = \sum_{a<x} p_X(a)$.
- Pois($\lambda$):
  * PMF: $p_X(k) = e^{-\lambda} \frac{\lambda^k}{k!}, k = 0, 1, 2, 3, \cdots$ and 0 otherwise.
  * CDF: $F_X(x) = \sum_{a<x} p_X(a)$.
- Expo($\lambda$):
  * PDF: $f_X(x) = \lambda e^{-\lambda x}$ for $x \in (0, \infty)$ and 0 otherwise.
  * CDF: $F_X(x) = \int_{-\infty}^{x} f_X(a)da$, which is equal to $\int_0^x f_X(a)da = 1 - e^{-\lambda x}$ for $x \in (0, \infty)$ and 0 otherwise.
- N($\mu, \sigma^2$):
  * PDF: $f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ for $x \in (-\infty, \infty)$.
  * CDF: $F_X(x) = \int_{-\infty}^{x} f_X(a)da$.

- Write down the properties of the aforementioned distributions plus uniform distribution. **(Need to know how to apply these properties to solve problems).**

**It is helpful to memorize these properties. You can use them directly during the exam by quoting these properties in your own words when solving exercises.**

- Bern(p): sum of independent Bern(p)-distributed r.v.'s follows Bin(n,p); coin tossing (either head or tail) and indicator r.v. follow Bern.
  Let $X \sim$Bern(p)

$$E(X) = p, V(X) = p(1-p).$$

- Bin(n,p): sum of independent Bern(p)-distributed r.v.'s follows Bin(n,p); Bin(1,p) is Bern(p); X forllows Bin(n,p) then $n - X$ follows Bin(n,1-p).
  Let $X \sim$Bin(n,p)

$$E(X) = np, V(X) = np(1-p).$$

- Let $X \sim$NBin(r,p)

$$E(X) = r(1-p)/p, V(X) = r(1-p)/p^2.$$

- Let $X \sim$HGeom(w,b,n)

$$E(X) = n\frac{w}{w+b}.$$

- Let $X \sim$NHGeom(w,b,r)

$$E(X) = r\frac{b}{w+1}.$$

- Pois($\lambda$): sum of independent Poisson-distributed r.v.'s still follows Poisson; Poisson diagram.
  Let $X \sim$Pois($\lambda$)

$$E(X) = \lambda, V(X) = \lambda.$$

- Expo($\lambda$): min of independent Expo-distributed r.v.'s still follows Expo; memoryless property.
  Let $X \sim$Expo($\lambda$)

$$E(X) = 1/\lambda, V(X) = 1/\lambda^2.$$

- N($\mu, \sigma^2$): sum of independent Normal-distributed r.v.'s still follows Normal; if $X$ follows $N(\mu, \sigma^2)$ then $(X-\mu)/\sigma$ follows standard normal N(0,1); symmetric.
  Let $X \sim$N($\mu, \sigma^2$)

$$E(X) = \mu, V(X) = \sigma^2.$$

- Unif: universality; Unif(a,b)-distributed $U$ conditional on $U \in (c,d) \subseteq (a,b)$ follows Unif(c,d).
  Let $X \sim$Unif(a,b)

$$E(X) = \frac{a+b}{2}, V(X) = \frac{(b-a)^2}{12}.$$