DS-GA 1002 - Homework 0

Eric Niblock

September 2nd, 2020

1. **(Sets) We will use set theory to define probability spaces. Are these statements true or false? Provide a proof if they are true (you can use Venn diagrams to gain intuition, but also write down a formal proof), or a counterexample if they are false.**

   A partition of a set Ω is a collection of sets $S_1, ..., S_n$ such that $\Omega = \cup_i S_i$ and $S_i \cap S_j = \emptyset$ for $i \neq j$

   (a) **If $S_1, ..., S_n$ is a partition of Ω, then for any subset $A \subseteq \Omega$, $S_1 \cap A, ..., S_n \cap A$ is a partition of A.**

   Proof by contradiction. Assume $S_1, ..., S_n$ is a partition of Ω and that for some $A \subseteq \Omega$, $S_1 \cap A, ..., S_n \cap A$ is not a partition of A. This implies that either,

$$A \neq \cup_i (S_1 \cap A, ..., S_n \cap A) \tag{1}$$

   Or that,

$$(S_i \cap A) \cap (S_j \cap A) \neq \emptyset \text{ for } i \neq j \tag{2}$$

   If we take case (1), this implies that there exists some $a \in A$ such that $a \notin \cup_i S_i$, since we know by the distributivity of intersections across unions that (1) can be written as, $A \neq \cup_i (S_1 \cap A, ..., S_n \cap A) = A \cap (\cup_i S_i)$. However, since $A \subseteq \Omega$, we know that $a \in \Omega$ and, by definition, $a \in \cup_i S_i$. Contradiction.

   If we take case (2), given that intersection is distributive over itself, we rewrite this condition as $(S_i \cap S_j) \cap A \neq \emptyset$ for $i \neq j$. However, since $S_1, ..., S_n$ is a partition of Ω, we know that $(S_i \cap S_j) = \emptyset$ for $i \neq j$. It follows that $\emptyset \cap A = \emptyset$. Contradiction.

   Therefore, it must be true that if $S_1, ..., S_n$ is a partition of Ω, then for any subset $A \subseteq \Omega$, $S_1 \cap A, ..., S_n \cap A$ is a partition of A.

(b) For any sets A and B, $A^C \cup B^C = (A \cup B)^C$.

False by example. Take $A, B \subseteq \Omega$ with $A = \{1\}$, $B = \{2\}$, and $\Omega = \{1, 2, 3\}$. It then follows that $A^C = \{2, 3\}$, $B^C = \{1, 3\}$, and $A^C \cup B^C = \{1, 2, 3\}$. On the other hand, $(A \cup B) = \{1, 2\}$, and therefore $(A \cup B)^C = \{3\}$. Thus, $A^C \cup B^C \neq (A \cup B)^C$.

(c) For any sets A, B, and C, $(A \cup B) \cap C = A \cup (B \cap C)$.

False by example. Take $A = \{1\}$, $B = \{2, 3\}$, and $C = \{3, 4\}$. Then $(A \cup B) = \{1, 2, 3\}$ and $(A \cup B) \cap C = \{3\}$. Furthermore, $(B \cap C) = \{3\}$, and $A \cup (B \cap C) = \{1, 3\}$. Therefore, $(A \cup B) \cap C \neq A \cup (B \cap C)$

2. (Series) We will need series to compute probabilities and expectations related to discrete quantities.

(a) Assuming $r \neq 1$, derive a simple expression for

$$S_n = \sum_{i=m}^{n} r^i \tag{3}$$

as a function of r, m and n, and prove that it holds. Assume m and n are positive integers with $m \leq n$.

We can express S_n as

$$S_n = r^m(1 + r + ... + r^{n-m}) = r^m \beta \tag{4}$$

Where we have defined β in (5). Multiplying by r yields the additional formula (6),

$$\beta = 1 + r + ... + r^{n-m} \tag{5}$$

$$r\beta = r + r^2 + ... + r^{n-m} + r^{n-m+1} \tag{6}$$

Subtracting (6) from (5) yields,

$$(1 - r)\beta = 1 - r^{n-m+1} \tag{7}$$

Solving produces,

$$\beta = \frac{1 - r^{n-m+1}}{1 - r} \tag{8}$$

And the final solution becomes,

$$S_n = \sum_{i=m}^{n} r^i = r^m \left(\frac{1 - r^{n-m+1}}{1 - r} \right) \tag{9}$$

We can prove that (9) is the correct closed formula for (3) by proceeding with induction on n, allowing m to be constrained only by the condition. First, we show that the base case holds when $(m, n) = (1, 1)$,

$$S_1 = \sum_{i=1}^{1} r^i = r \tag{10}$$

$$S_1 = r\left(\frac{1 - r^{1-1+1}}{1 - r}\right) = r\left(\frac{1 - r}{1 - r}\right) = r \tag{11}$$

With the base case satisfied we take the hypothesis to be (3), with $n = k$ and $m = m$.

$$S_k = \sum_{i=m}^{k} r^i = r^m\left(\frac{1 - r^{k-m+1}}{1 - r}\right) \tag{12}$$

Now, the closed form solution with $n = k + 1$ and $m = m$ is,

$$S_{k+1} = \sum_{i=m}^{k+1} r^i = r^m(1 + r + ... + r^{k+1-m}) \tag{13}$$

In the same process as before, we define β, multiply by r, subtract, and solve for β,

$$\beta = 1 + r + ... + r^{k+1-m} \tag{14}$$

$$r\beta = r + r^2 + ... + r^{k+2-m} \tag{15}$$

$$\beta = \frac{1 - r^{k+2-m}}{1 - r} \tag{16}$$

The final result becomes,

$$S_{k+1} = \sum_{i=m}^{k+1} r^i = r^m\left(\frac{1 - r^{(k+1)-m+1}}{1 - r}\right) \tag{17}$$

The induction is therefore complete, and the solution provided in (9) is proven.

(b) Under what condition on r does the infinite series

$$\sum_{i=m}^{\infty} r^i = \lim_{n\to\infty} S_n \tag{18}$$

converge (where again m is a positive integer)?

First, we begin by simplifying the limit given in (18),

$$\lim_{n\to\infty} S_n = \lim_{n\to\infty} r^m \left(\frac{1-r^{n-m+1}}{1-r}\right) = \frac{r^m}{1-r}(1 - \lim_{n\to\infty} r^{n-m+1}) \tag{19}$$

Now, just by analyzing the contents of the limit, we see that r^{n-m+1} will converge if $-1 < r < 1$ when $n \to \infty$. Observe that we cannot have $r = 1$ because this would create a zero in the denominator. Therefore, the infinite series in (18) converges when $-1 < r < 1$.

(c) Use induction to prove the identity

$$\sum_{i=1}^{n} i = \frac{n(n+1)}{2} \tag{20}$$

where n is a non-negative integer greater than 1.

First, we must show that the base case is satisfied. We take this to be when $n = 2$.

$$\sum_{i=1}^{2} i = \frac{2(2+1)}{2} = 3 \tag{21}$$

Both the left and right sides of (21) provide the same result - the base case is therefore satisfied. Next, we set our hypothesis to be equation (20), though with $n = k$,

$$\sum_{i=1}^{k} i = \frac{k(k+1)}{2} \tag{22}$$

And finally, we show the induction by using $n = k + 1$,

$$\sum_{i=1}^{k+1} i = \frac{(k+1)(k+2)}{2} = \frac{k(k+1)}{2} + (k+1) \tag{23}$$

This proves (20) through induction, because the result of (23) is $k + 1$ greater than (22), as expected.

3. **(Derivatives) We will use derivatives to define probability density functions. The derivative of a differentiable function f is defined as**

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \tag{24}$$

(a) **Briefly explain why the derivative of a function can be interpreted as an *instantaneous rate of change.***

The formula for a derivative of a function is remarkably similar to the formula for the slope of a line between two points. The main difference is that the derivative yields the rate of change at a single point along a function, whereas the slope of a line determines the average rate of change between two different points. A derivative accomplishes this task by shrinking the difference between two points to be that of infinitesimal magnitude. The notion of determining the rate of change at some precise point contributes the 'instantaneous' in *instantaneous rate of change*. We refer to both the derivative and the slope as 'rates of change' because they are measuring the change in some y-variable as a result of some change in x-variable.

(b) **Use the definition to derive the derivative of the function x^2.**

Employing the definition of the derivative when $f(x) = x^2$, we find,

$$\begin{aligned} f'(x) &= \lim_{h \to 0} \frac{(x+h)^2 - x^2}{h} = \lim_{h \to 0} \frac{2xh + h^2}{h} \\ &= \lim_{h \to 0} 2x + h = 2x \end{aligned} \tag{25}$$

(c) **We would like to approximate a differentiable function f at y using a linear function $L_y(x) = ax + b$. We set a and b so that f and L_y have the same value and the same derivative at y (i.e., $L_y(y) = f(y)$ and $L'_y(y) = f'(y)$). Give an expression for $L_y(x)$ in terms of y, $f(y)$, and $f'(y)$.**

Given the conditions, we have,

$$L_y(y) = ay + b = f(y) \tag{26}$$

And therefore,

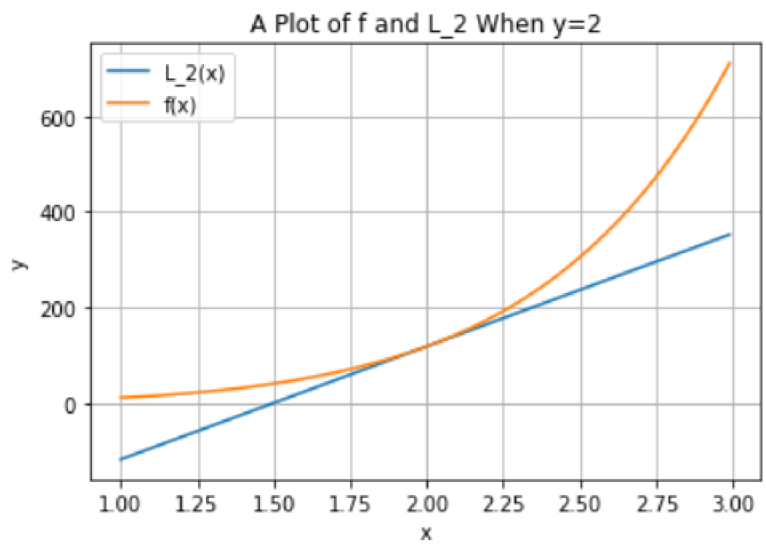
$$
\begin{aligned}
a &= f'(y) \\
b &= f(y) - f'(y)y
\end{aligned} \tag{27}
$$

So the complete equation becomes,

$$L_y(x) = f'(y)(x - y) + f(y) \tag{28}$$

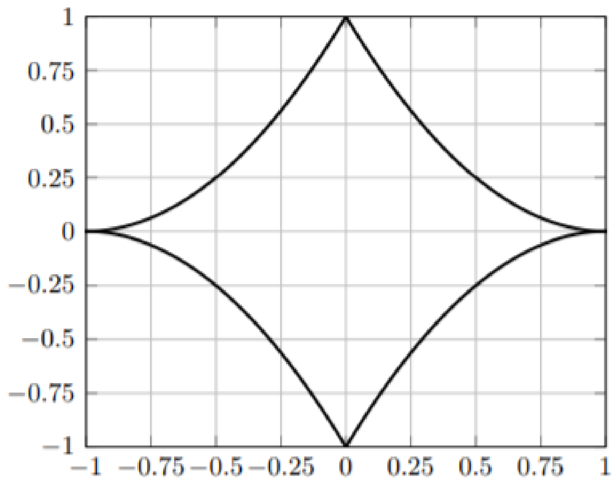
(d) Let $f(x) = 4x^2e^x$. Plot f and L_2 between 1 and 3.

Observe the following plot showing the relationship between f and L_2,



4. **(Integrals) We will use integrals to compute probabilities and expectations related to continuous quantities.**

(a) **Express the area of the following shape in terms of an integral and solve it. Each of the four bounding curves are graphs of quadratic functions. As depicted, the bounding curve includes the points $(0,1)$, $(1,0)$, and $(1/2,1/4)$, and is symmetric about the x and y axes.**



By inspection, the quadratic in the upper left corner is $f(x) = (x+1)^2$. Due to the symmetry of the figure, we can find the area of the region trapped beneath this curve (bounded by the two axes) and multiply by 4. The integral then reads,

$$4\int_{-1}^{0}(x+1)^2dx = \frac{4(x+1)^3}{3}\Big|_{-1}^{0} = \frac{4}{3} \tag{29}$$

(b) **Use change of variables to derive a closed-form expression for the function**

$$f(t) = \int_0^t \frac{x}{1+x^2}dx \tag{30}$$

We employ the following change of variables: $u = 1 + x^2$ and $du = 2xdx$. The integral then becomes,

$$f(t) = \int_1^{1+t^2} \frac{du}{2u} = \frac{1}{2}ln(u)\Big|_1^{1+t^2} = \frac{1}{2}ln(1+t^2) \tag{31}$$