AIM Preliminary Exam: Probability and Discrete Mathematics

August 20, 2021

There are five (5) problems in this examination.

There should be sufficient room in this booklet for all your work. But if you use other sheets of paper, be sure to mark them clearly and staple them to the booklet.

Problem 1

For a nonnegative integer-valued random variable N, show that

$$E[N] = \sum_{i=1}^{\infty} P[N \geq i].$$

Solution

$$\sum_{i=1}^{\infty} P[N \geq i] = \sum_{k \geq i} P[N = k] = \sum_{k=1}^{\infty} \sum_{i=1}^{k} P[N = k] = \sum_{k=1}^{\infty} P[N = k] \sum_{i=1}^{k} 1 = \sum_{k=1}^{\infty} kP[N = k] = E[N].$$

Mathematical concepts: discrete random variables, expectation, interchanging summation (combinatorics)

Problem 2

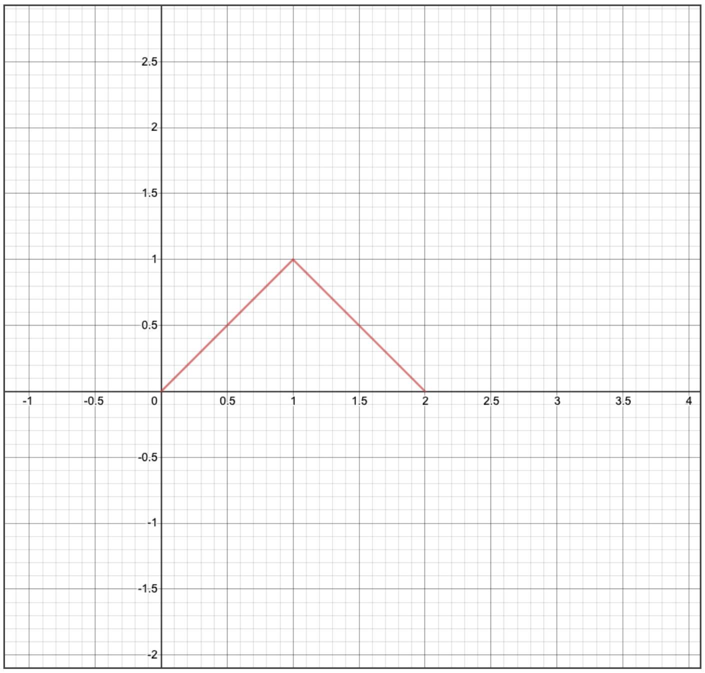
Suppose X, Y, Z are i.i.d. real-valued random variables uniformly-distributed on $[0, 1]$.

(a) Sketch the density $f_{X+Y}(x)$ on a Cartesian plane with f_{X+Y} vertical and x horizontal. Label algebraically, where appropriate. Also illustrate (somehow) the joint density for (X, Y) on a Cartesian plane with x horizontal and y vertical. Finally, also illustrate the density of $X + Y$ on this graph with x horizontal and y vertical.

Solution

Triangle: $(0, 0)$ to $(1, 1)$ to $(2, 0)$.

12/29/2020 Desmos | Graphing Calculator

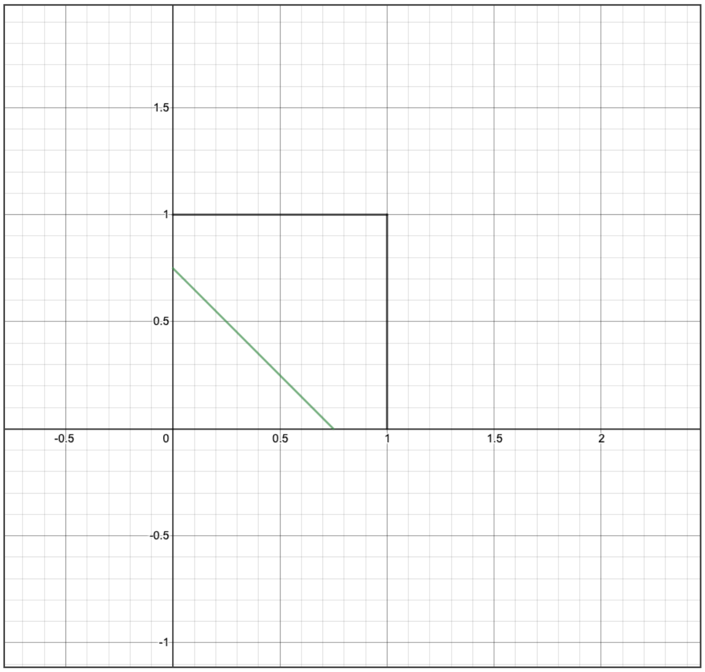


https://www.desmos.com/calculator 1/2

$$f_{X+Y}(x) = \begin{cases} x, & 0 \le x \le 1; \\ 2 - x, & 1 \le x \le 2; \\ 0, & \text{otherwise.} \end{cases}$$

The function is the value of the cross-sectional length of the unit square, measured perpendicular to the diagonal $y = x$, to the point $(x, 0)$.

12/29/2020 Desmos | Graphing Calculator



https://www.desmos.com/calculator 1/2

(b) Sketch the density $f_{X+Y+Z}(x)$ on a Cartesian plane with f_{X+Y+Z} vertical and x horizontal. Label algebraically, where appropriate.

Solution

These are done by convolution. For $0 \le x \le 1$, we have $f_{X+Y}(x) = x$ from above. So

$$f_{X+Y+Z}(x) = \iint_{u+v=x} f_{X+Y}(u) f_Z(v)\, du\, dv = \int_{u=0}^{x} f_{X+Y}(u) f_Z(x-u)\, du.$$

Note that $f_Z = \chi_{[0,1]}$. For the interval $0 \le x \le 1$, we get $\int_0^x u\, du = \boxed{\frac{x^2}{2}}$. By symmetry (under $X \Longleftrightarrow 1-X, \quad Y \Longleftrightarrow 1-Y, \quad Z \Longleftrightarrow 1-Z, \quad X+Y+Z \Longleftrightarrow 3-(X+Y+Z)$), for $2 \le x \le 3$, we get $\boxed{\frac{(3-x)^2}{2}}$.

The total probability so far is $2\int_0^1 \frac{x^2}{2}\, dx = \frac{1}{3}$. For $1 \le x \le 2$, the density takes the value 1/2 at each endpoint, is symmetrical, a parabola, and has integral $1 - 1/3 = 2/3$. So $f_{X+Y+Z}(x) = \boxed{\frac{3}{4} - \left(\frac{3}{2} - x\right)^2}$.

Alternatively, for the interval $1 \le x \le 2$, we have $x - 1 \le u \le x$ in order that the argument $x - u$ to f_Z

makes $f_Z = \chi_{[0,1]}$ take the value 1. We have

$$
\begin{aligned}
f_{X+Y+Z}(x) &= \int f_{X+Y}(u) f_Z(x-u)\,du \\
&= \int_{u=x-1}^{x} f_{X+Y}(u)\,du \\
&= \int_{u=x-1}^{1} f_{X+Y}(u)\,du + \int_{u=1}^{x} f_{X+Y}(u)\,du \\
&= \int_{u=x-1}^{1} u\,du + \int_{u=1}^{x} (2-u)\,du \\
&= \left[\frac{1}{2} - \frac{(x-1)^2}{2}\right] - \left[\frac{(2-x)^2}{2} - \frac{1}{2}\right] \\
&= \frac{3}{4} - \left(\frac{3}{2} - x\right)^2 .
\end{aligned}
$$

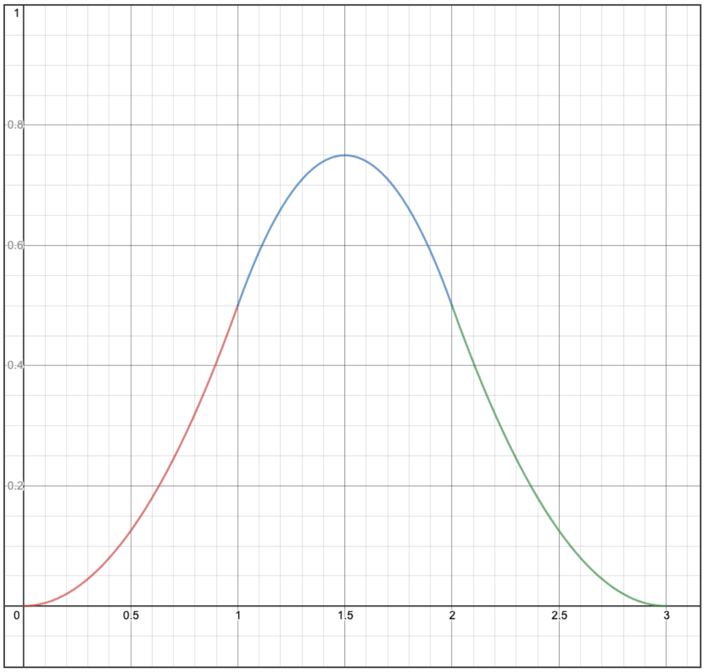
In the penultimate line, we can confirm the value $1/2$ at $x = 1$ or $x = 2$, a parabola, so symmetric about $x = 3/2$, with leading term $-x^2$. Plugging in $x = 3/2$, we get the value $3/4$.

So

$$
f_{X+Y+Z}(x) = \begin{cases} \frac{x^2}{2!}, & 0 \le x \le 1; \\ \frac{3}{4} - \left(\frac{3}{2} - x\right)^2, & 1 \le x \le 2; \\ \frac{(3-x)^2}{2!}, & 2 \le x \le 3; \\ 0, & \text{otherwise.} \end{cases}
$$

Three parabolas, smiling, frowning, smiling, symmetric.

12/29/2020 Desmos | Graphing Calculator



https://www.desmos.com/calculator 1/2

Mathematical concepts: joint distributions, continuous random variables, sum of random variables

Problem 3

(a) Use the binomial theorem to prove that $2^n = \sum_{k=0}^{n}(-1)^k\binom{n}{k}3^{n-k}$.

Solution

$$2^n = (3-1)^n = \sum_{k=0}^{n}(-1)^k\binom{n}{k}3^{n-k}.$$

(b) Use combinatorial reasoning to prove, for $n \geq 3$, that

$$\binom{n}{k} - \binom{n-3}{k} = \binom{n-1}{k-1} + \binom{n-2}{k-1} + \binom{n-3}{k-1}.$$

Solution

Suppose a set S of size $n \geq 3$ has three distinguished objects, a, b, c. The left hand side counts the total number of k-sets less the number of k-sets that omit all of a, b, c. The right hand side counts the disjoint collections of the number of sets that have a and a $(k-1)$-set of $S - a$, the number of sets having b, no a, and a $(k-1)$-set of $S \setminus \{a, b\}$, and the sets having c, no a, and no b, and a $(k-1)$-set of $S \setminus \{a, b, c\}$. Both are the number of k-sets of S having a, b, or c.

Mathematical concepts: Binomial theorem.

Problem 4

Prove that, if we choose five points in a unit square, some pair of the points are at distance at most $1/\sqrt{2}$.

Solution

Partition the square into a 2×2 array of four subsquares of side $1/2$. Then, by the pigeonhole principle, some subsquare has two points. The diameter of that subsquare is $1/\sqrt{2}$.

Mathematical concepts: Pigeonhole principle.

Problem 5

Let G be an undirected graph with distinct edge weights. Let S be a subset of the vertex set V that is neither empty nor V. Let $e = (v, w)$ be the minimum-cost edge with $v \in S$ and $w \in V \setminus S$.

Prove that every minimum spanning tree contains e.

Solution

Suppose T is a spanning tree without e; we need to show that T does not have minimum cost. To do this, we'll find another edge, $e' \in T$, such that exchanging e with e' maintains the tree property but lowers the cost. That is, e' costs more than e.

Let P be a path in T from v to w. Let $e' = (v', w')$ be the first edge along P with $v' \in S$ and $w' \in V \setminus S$. Such an edge exists because $v \in S$ and $w \in V \setminus S$.

Then $T' = T - e' + e$ is a spanning tree, because any path in T that uses e' could instead use $P - e' + e$. Similarly, the graph T' is a tree, because if T' has a cycle C (which must involve e), then $C - e \cup P$ is a cycle in T—contradiction. The cost of T' is less than the cost of T because e has minimum cost among edges from S to $V \setminus S$ (and the costs are distinct).

Mathematical concepts: Minimum spanning tree, greedy algorithm