151[1] *A box contains 13 sound grapefruit and three that are rotten. You pick four at random.*

(a) Describe the sample space.

(b) Let X be the number of sound grapefruit you pick. Find $f_X(x)$ and $\mathbf{E}(X)$.

(a.) This is an example of a hypergeometric random variable because the four grapefruit are drawn without replacement. Thus the sample space Ω is all combinations of the $13+3$ grapefruit taken four at a time. In other words, it is the collection of subsets of size four from all 16 grapefruit. Thus $|\Omega| = \binom{16}{4}$.

(b.) If X is the number of sound grapefruit among those you pick, then $D = X(\Omega) = \{1, 2, 3, 4\}$ are the values taken by X. Since the number of rotten grapefruit in my selection is at most the number of rotten grapefruit in the box, $4 - x \le 3$, we must have $x \ge 1$. The number of combinations with x sound grapefruit and $4 - x$ rotten ones is the number of ways of choosing a subset of size x from 13 times the number of ways of choosing the remaining $4 - x$ from the 3 rotten ones. Thus if $x \in D$,

$$f_X(x) = \frac{\binom{13}{x}\binom{3}{4-x}}{\binom{16}{4}},$$

and $f_X(x) = 0$ if $x \notin D$. Thus

$$f_X(1) = \frac{\binom{13}{1}\binom{3}{3}}{\binom{16}{4}} = \frac{1}{140}, \qquad f_X(2) = \frac{\binom{13}{2}\binom{3}{2}}{\binom{16}{4}} = \frac{18}{140},$$
$$f_X(3) = \frac{\binom{13}{3}\binom{3}{1}}{\binom{16}{4}} = \frac{66}{140}, \qquad f_X(4) = \frac{\binom{13}{4}\binom{3}{0}}{\binom{16}{4}} = \frac{55}{140}.$$

It follows that the expectation

$$\begin{aligned}\mathbf{E}(X) &= \sum_{i=1}^{4} i f_X(i) \\ &= \frac{1\cdot 1 + 2\cdot 18 + 3\cdot 66 + 4\cdot 55}{140} = \frac{1 + 36 + 198 + 220}{140} = \frac{455}{140} = \frac{13}{4}.\end{aligned}$$

A. *Roll four dice. Let X be the number of sixes. Find Ω, D, $f_X(x)$, $F_X(x)$ and $\mathbf{E}(X)$. Graph $f_X(x)$ and $F_X(x)$.*

$X \sim \text{binom}(n, p)$ is distributed according to a binomial distribution. The $n = 4$ rolls of the dice are assumed independent and the probability of success, that is of rolling a six, is $p = \mathbf{P}(S) = \frac{1}{6}$. The sample space is the record of four rolls

$$\Omega = \{(x_1, x_2, x_3, x_4) : \ x_i \in \{S, F\} \text{ for all } i.\}.$$

Alternately, the sample space could be the four-tuple of numbers rolled. $D = X(\Omega\} = \{0, 1, 2, 3, 4\}$ is the set of possible values of the random variable. For $x \in D$, the probability of x sixes is the number of ways to get x sixes in four rolls times the probability of getting a particular sequence of x sixes and $4 - x$ non-sixes

$$f_X(x) = \binom{4}{x} p^x \, q^{4-x}$$

x	$f_X(x)$	$F_X(x)$
0	$\binom{4}{0}\left(\frac{1}{6}\right)^0\left(\frac{5}{6}\right)^4 = \frac{5^4}{6^4} = \frac{625}{1296}$	$\frac{625}{1296}$
1	$\binom{4}{1}\left(\frac{1}{6}\right)^1\left(\frac{5}{6}\right)^3 = \frac{4\cdot 5^3}{6^4} = \frac{500}{1296}$	$\frac{625+500}{1296} = \frac{1125}{1296}$
2	$\binom{4}{2}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^2 = \frac{6\cdot 5^2}{6^4} = \frac{150}{1296}$	$\frac{625+500+150}{1296} = \frac{1275}{1296}$
3	$\binom{4}{3}\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^1 = \frac{4\cdot 5}{6^4} = \frac{20}{1296}$	$\frac{625+500+150+20}{1296} = \frac{1295}{1296}$
4	$\binom{4}{4}\left(\frac{1}{6}\right)^4\left(\frac{5}{6}\right)^0 = \frac{1}{6^4} = \frac{1}{1296}$	$\frac{625+500+150+20+1}{1296} = 1$

Figure 1: Table of p.d.f and c.d.f for $x \in D$.

and $f_X(x) = 0$ if $x \notin D$. Values are given in Figure 1. The cumulative distribution function for $x \in D$ is given by

$$F_X(x) = \sum_{j \in D \text{ and } j \le x} f_X(j) = \sum_{j=0}^{x} f_X(j).$$

Note that for any real x, $F_X(x) = 0$ if $x < 0$, $F_X(x) = F_X(\lfloor x \rfloor)$ if $0 \le x \le 4$ and $F_X(x) = 1$ if $x > 4$. ($\lfloor x \rfloor$ is the greatest integer function of x.) The expectation is

$$\begin{aligned}\mathbf{E}(X) &= \sum_{i=0}^{4} i f_X(i) \\ &= \frac{0\cdot 625 + 1\cdot 500 + 2\cdot 150 + 3\cdot 20 + 4\cdot 1}{1296} = \frac{500+300+60+4}{1296} = \frac{864}{1296} = \frac{2}{3}.\end{aligned}$$

Graphing gives

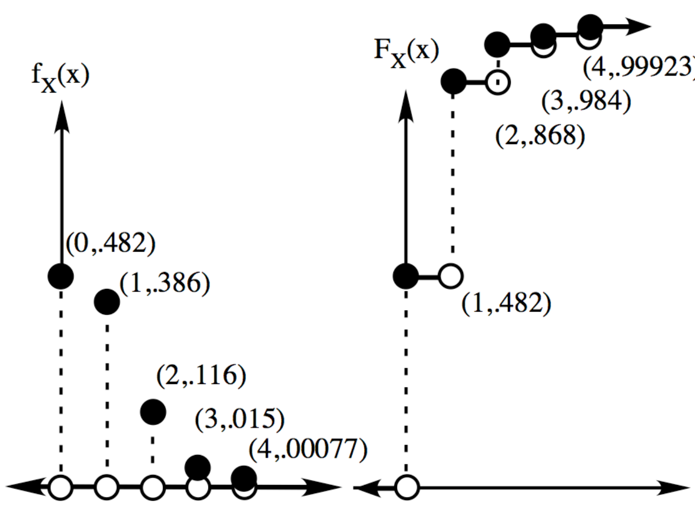

Figure 2: Graphs of $f_X(x)$ and $F_X(x)$.

B. *let* $D = \{1, 2, 3, \ldots, n\}$. *Let* $f(x) = cx$ *for* $x \in D$ *and* $f(x) = 0$ *for* $x \notin D$. *Find* c *to make* $f(x)$ *a p.d.f. Find* $F_X(x)$ *and* $\mathbf{E}(X)$.

To be a probability density function, we need $f(x) \ge 0$ so $c > 0$ must hold. Also the total probability satisfies

$$1 = \sum_{x \in D} f(x) = \sum_{i=1}^{n} ci = \frac{cn(n+1)}{2},$$

using the formula for the sum of the first n numbers. Hence

$$c = \frac{2}{n(n+1)}.$$

Using the same formula we get the cumulative distribution function. For $x \in D$,

$$F_X(x) = \sum_{j \,\in\, D \text{ and } j \,\le\, x} f(j) = \sum_{j=1}^{x} c\, j = \frac{x(x+1)}{n(n+1)}.$$

Thus for any real x, $F_X(x) = 0$ if $x < 1$, $F_X(x) = 1$ if $x \ge n$ and $F_X(x) = F_X(\lfloor x \rfloor)$ if $1 \le x \le n$, where $\lfloor x \rfloor$ denotes the greatest integer function.

Using the formula for the sum of the first n squares, the expectation is

$$\mathbf{E}(X) = \sum_{x \in D} x\, f(x) = \sum_{i=1}^{n} c\, i^2 = \frac{2}{n(n+1)} \cdot \frac{n(n+1)(2n+1)}{6} = \frac{2n+1}{3}.$$