5.2 Linearity of Expectation:

The most important property of expectation is linearity: the expected sum of r.v.s is the sum of the individual expected values.

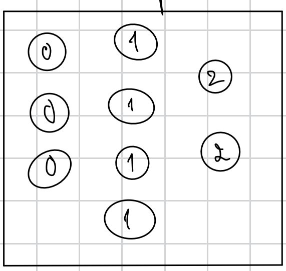
Theorem 5.2.1 (Linearity of Expectation)

For any r.v.s X, Y and constant c:

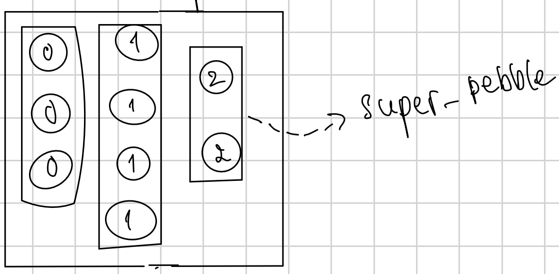
- $E(X+Y) = E(X) + E(Y)$,
- $E(cX) = c.E(X)$

There are 2 ways of calculating averages: grouped and ungrouped

Take examlep of r.v X assigning values to pebbles



X assign a number to each pebble in sample space



Group pebbles by values, 9 pebbles now becomes 3 super-pebbles. The weight of a supper-pebble is the sum of the weights of the constituent pebbles.

- Grouped averages: the process of creating super-pebbles whose weight, $P(X=x)$, is the total weight of the constituent pebbles. For example: $\underbrace{\frac{3}{9}}_{\text{weights}}.0 + \frac{4}{9}.1 + \frac{2}{9}.2$

This definition helps us works with distribution of X directly without returning to the sample space. However, if there is another r.v Y on the same sample space who creates super-pebbles with different weights $P(Y=y)$, it will be difficult to calculate the combination of $\sum_x x P(X=x)$ and $\sum_y y P(Y=y)$.

That is where ungrouped averages comes in.

- Ungrouped averages: The advantage of this definition is that it breaks down the sample space into the smallest units. For example: $\underbrace{\frac{1}{9}}_{\text{smallest unit}}(0+0+\ldots+1+1+\ldots+2)$

For more generality:

- $E(X) = \sum_s X(s) \, P(\{s\})$

value that X assigns to pebble s

weight of pebble s

- $E(Y) = \sum_s Y(s) \, P(\{s\})$

Now we can combine $\sum_s X(s) P(\{s\})$ and $\sum_s Y(s) P(\{s\})$, which gives $E(X+Y)$

So: $E(X) + E(Y) = \sum_s X(s) P(\{s\}) + \sum_s Y(s) P(\{s\})$

$$= \sum_s (X+Y)(s) \, P(\{s\})$$

$$= E(X+Y)$$

Example 5.2.3 (Binomial expectation) Find $E(X)$ of $X \sim Bin(n,p)$

- By definition of expectation: $E(X) = \sum_{k=0}^{n} k \cdot P(X=k) = \sum_{k=0}^{n} k \cdot \binom{n}{k} p^k q^{n-k}$
- Linearity of expectation can provide a much shorter path to the same result:
  - Write X as sum of n i.i.d Bern(p) r.v.s: $X = I_1 + \dots + I_n$,

    where each I_j has expectation $E(I_j) = 1p + 0q$
    $= p$
  - Then by linearity: $E(X) = E(I_1) + \dots + E(I_j) = n \cdot p$

Example 5.2.4 (Hypergeometric expectation) Find $E(X)$ of $X \sim HGeom(w,b,n)$

- We can interpret the problem as X is the number of white balls in sample size n drawn from an urn consists of w white balls and b black balls without replacement
- We can write X as sum of Bernoulli r.v.s: $X = I_1 + \dots + I_n$,

$$\text{where } I_j = \begin{cases} 1 & \text{if } j^{th} \text{ ball is white} \\ 0 & \text{otherwise} \end{cases}$$

By symmetry, $I_j \sim Bern(p)$ with $p = \frac{w}{w+b}$ (since unconditionally the j^{th} ball drawn is equally likely to be any of the balls)

- However, since its sampling without replacement, the I_j are not independent. However, linearity still holds for dependent r.v.s:

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