1. Explain the concept of Expected Value. (Use some mathematical goodness).
   * Expected value is the probability-weighted average of all possible values, e.g. in a 6-sided die, the expected value is (1+2+3+4+5+6)/6 = 3.5 cause all sides have similar probability, hence expected value is the average value of a random variable over many experiments.
   * If the probability of each item varies like the following example (number of workouts per week):

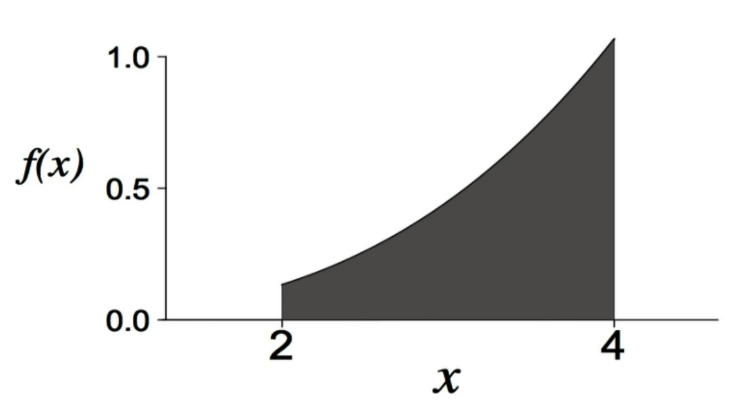
|  |  |
| --- | --- |
| Number of Workouts (X) | Probability |
| 0 | 0.1 |
| 1 | 0.15 |
| 2 | 0.4 |
| 3 | 0.25 |
| 4 | 0.1 |

Expected Value = Value \* Probability,

E(x) = ΣX\*p(x) = (0.1\*0 + 0.15\*1 + 0.4\*2 + 0.25\*3 + 0.1\*4) = 2.1

1. Write down the equations for mean(m), variance(var) and standard deviation(stddev) of a population vs. a sample.
2. What is the average for a continues variable. With an example.

For a continues distribution the mean value is equal to the expected value of Probability Distribution Factor:



where *f(x)* is the Probability Distribution Factor

As for variance:

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Age | SEX | BMI | BP | S1 | S2 | S3 | S4 | S5 | S6 | target |
| mean | 48.52 | 1.47 | 26.38 | 94.65 | 189.14 | 115.44 | 49.79 | 4.07 | 4.64 | 91.26 | 152.13 |
| variance | 171.85 | 0.25 | 19.52 | 191.30 | 1,197.72 | 924.96 | 167.29 | 1.67 | 0.27 | 132.17 | 5,943.33 |
| standard dev | 13.11 | 0.50 | 4.42 | 13.83 | 34.61 | 30.41 | 12.93 | 1.29 | 0.52 | 11.50 | 77.09 |
| covariance | 189.88 | 1.66 | 199.75 | 470.75 | 565.68 | 408.09 | -393.66 | 42.82 | 22.79 | 338.99 | 5,943.33 |
| correlation | 0.19 | 0.04 | 0.59 | 0.44 | 0.21 | 0.17 | -0.39 | 0.43 | 0.57 | 0.38 | 1.00 |

There is a positive relationship between BMI and Target, where both variables move in the same direction

There is a negative relationship between S3 and Target, as S3 increases Target value