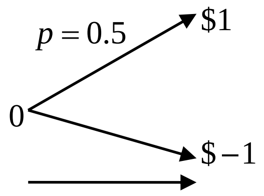
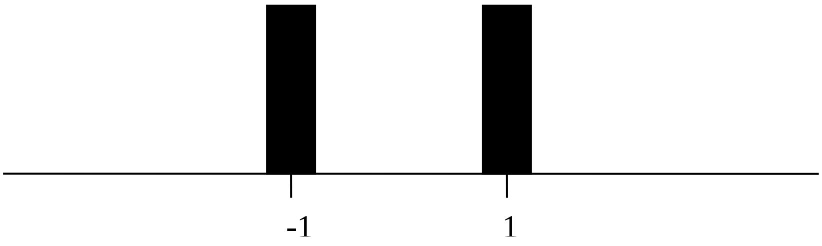
Identical Independent Random Variables

Imagine a random event that involves the tossing of a single coin. We have two outcomes, heads or tails. Let us define a random variable x_1 on the outcomes that takes the value of plus one dollar if we obtain a head and minus one dollar if we obtain a tail.

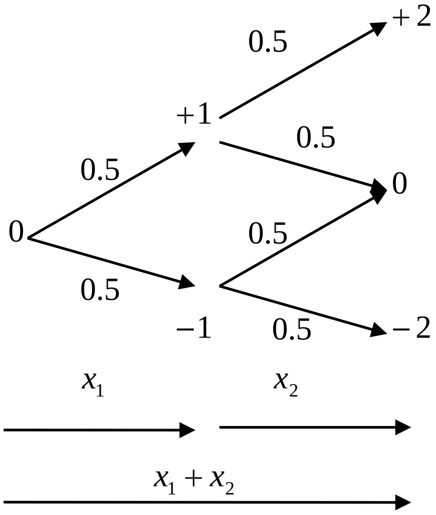


The distribution of our profit and loss is then as follows. The mean, μ of the distribution is zero and variance, σ^2 is one.

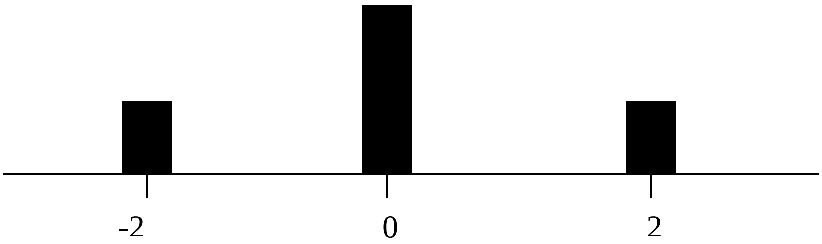


- $\mu = 0$
- $\sigma^2 = 0.5(1-0)^2 + 0.5(-1-0)^2 = 1.0$

If we play the original game twice then our new game effectively involves two coin tosses.

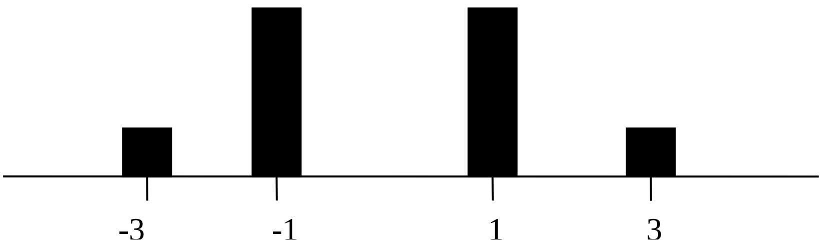


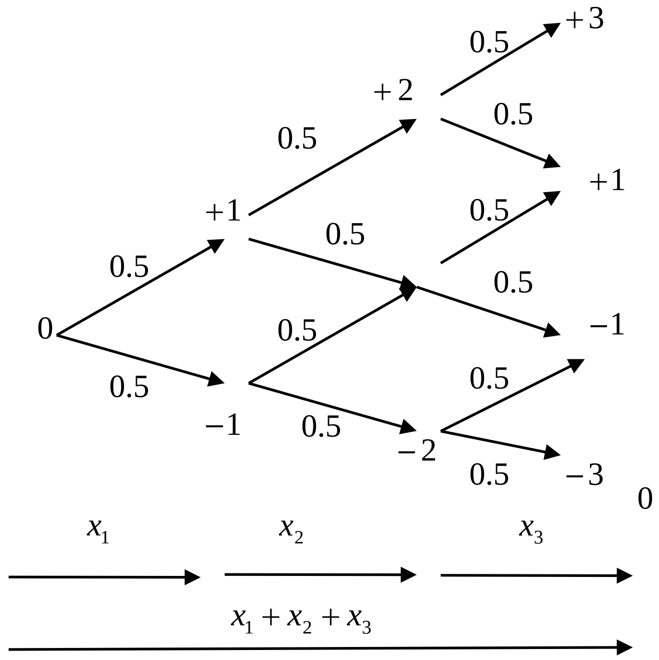
If we define $X_2 = x_1 + x_2$ then the distribution of our profit and loss is as follows. The mean of the distribution is zero and the variance is two.



- $\mu = 0$
- $\sigma^2 = 0.25(2-0)^2 + 0.25(-2-0)^2 = 2.0$

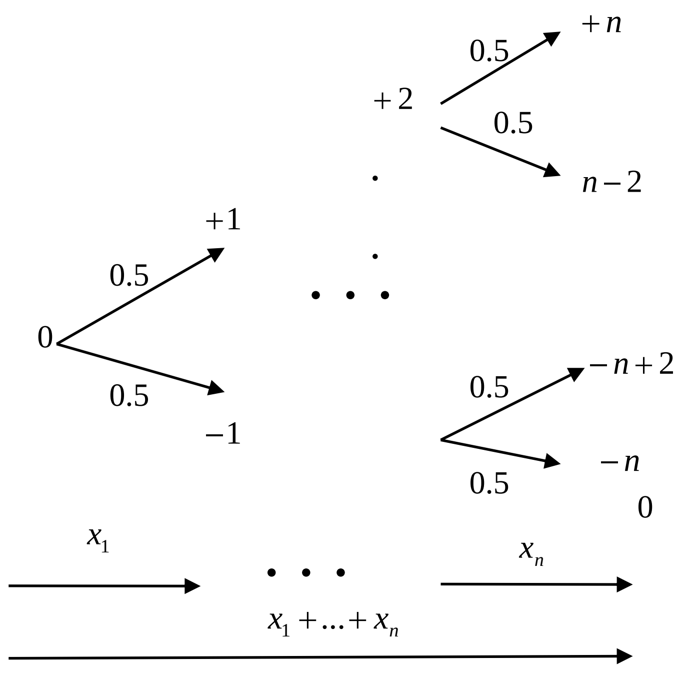
Let us go one-step further and look at the event obtained by summing three of the original events. $X_3 = x_1 + x_2 + x_3$ We get the following distribution, whose mean is zero and whose variance is three.





- $\mu = 0$
- $\sigma^2 = 0.125(3-0)^2 + 0.125(-3-0)^2 + 0.375(1-0)^2 + 0.375(-1-0)^2 = 3.0$

Generalising, if we perform n identical tosses of a coin and define n identical random variables $x_1, x_2, x_3, \ldots, x_n$ each will also have mean zero, and variance, σ^2 of one. By summing them we obtain the random variable $X_n = x_1 + \ldots + x_n$ which is distributed with mean zero and variance n.



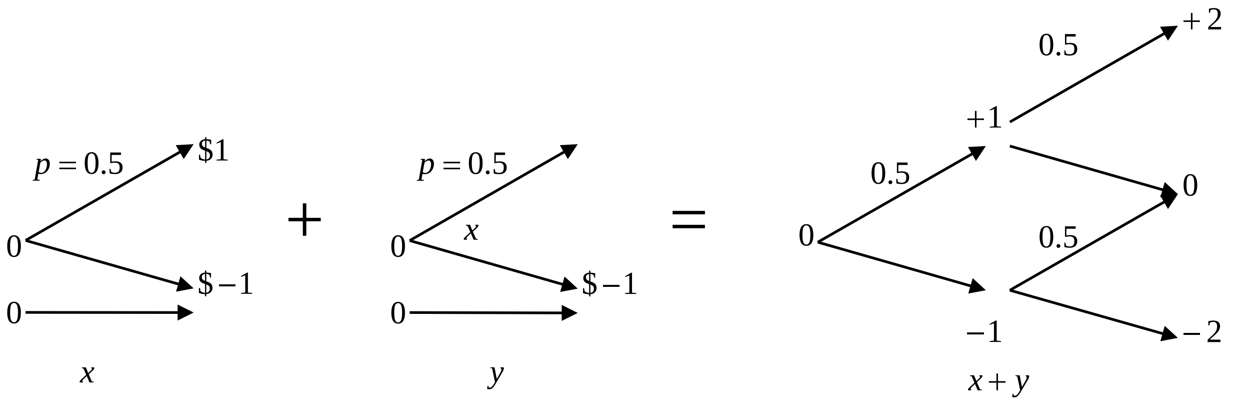
From a proof of why the sum of n independent, identically distributed random variables with mean μ and variance σ^2 is a random variable with mean $n\mu$ and variance $n\sigma^2$ see below

Proofs

Expectation of sum of N I.I.D random variables

$$E[X + X + X + X+ \ldots\ldots + X] = NE[X]$$

The following figure shows the general approach



We have random variable x with sample space $\{x_1, x_2\} = \{1,0\}$ and a second identically distributed random variable $y$with sample space$\{y_1, y_2\} = \{1,0\}$. The sample space of the joint distribution $x + y$ is given by the set of pairs

$$\{x_1, y_1\}, \{x_1, y_2\}$$

$$\{x_2, y_1\}, \{x_2, y_2\}$$

The expectation of the sum of the variables is then given by

$$\sum_{i=1}^{2}\sum_{j=1}^{2}(x_i + y_j)p(x_i, y_j)$$

Multiplying out

$$\sum_{i=1}^{2}\sum_{j=1}^{2}x_i\, p(x_i, y_j) + \sum_{i=1}^{2}\sum_{j=1}^{2}y_j\, p(x_i, y_j)$$

Noting that $\sum_{j=1}^{n} p(x_i\,, y_j) = p(x_i)$ and $\sum_{i=1}^{m} p(x_i\,, y_j) = p(y_j)$

$$\sum_{i=1}^{m}x_i p(x_i) + \sum_{j=1}^{n}x_j p(y_j)$$

Therefore we can note that

$$\boldsymbol{E}[\boldsymbol{X}+\boldsymbol{Y}] = \boldsymbol{E}[\boldsymbol{X}] + \boldsymbol{E}[\boldsymbol{Y}]$$

Generalising

If x is a random variable with sample space $\{x_1, x_2, \ldots\ldots, x_m\}$ and y is an independent random variable with sample space $\{y_1, y_2, \ldots\ldots, y_n\}$then the sample space of the joint distribution will be given by a set of pairs

$\{x_1, y_1\}, \{x_1, y_2\}, \ldots\ldots, \{x_1, y_n\}$

$\{x_2, y_1\}, \{x_2, y_2\}, \ldots\ldots, \{x_2, y_n\}$

.

.

$\{x_m, y_1\}, \{x_m, y_2\}, \ldots\ldots, \{x_m, y_n\}$

The expectation of the sum of the two variables in then given by

$\sum_{i=1}^{m} \sum_{j=1}^{n} (x_i + y_j) p(x_i, y_j)$

Multiplying out we get

$$\sum_{i=1}^{m} \sum_{j=1}^{n} x_i\, p(x_i, y_j) + \sum_{i=1}^{m} \sum_{j=1}^{n} y_j\, p(x_i, y_j)$$

Noting that $\sum_{j=1}^{n} p(x_i, y_j) = p(x_i)$ and $\sum_{i=1}^{m} p(x_i, y_j) = p(y_j)$

$$\sum_{i=1}^{m} x_i p(x_i) + \sum_{j=1}^{n} x_j p(y_j)$$

Therefore we can note that

$$E[X+Y] = E[X] + E[Y]$$

Furthermore

$$E[X+X] = E[X] + E[X] = 2E[X]$$

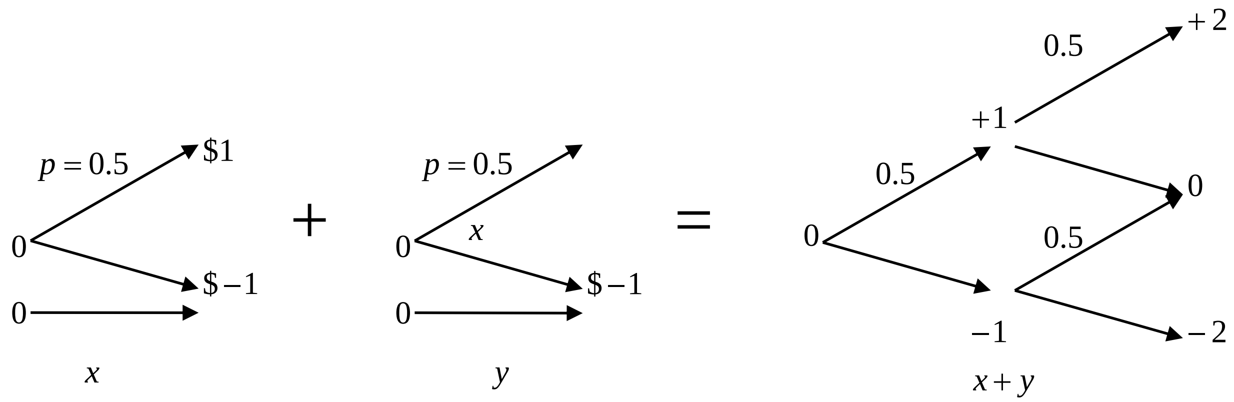
And

$$E[X + X + X + X+\dots\dots+X] = NE[X]$$

Variance of sum of I.I.D random variables

$$Var[x + y] = Var[x] + Var[y]$$

The following diagram shows the general approach.



We have random variable x with sample space $\{x_1, x_2\} = \{1,0\}$and another identically distributed random variable y with sample space$\{y_1, y_2\} = \{1,0\}$. The sample space of the joint distribution $x + y$is given by the set of pairs

$$\{x_1, y_1\}, \{x_1, y_2\}$$

$$\{x_2, y_1\}, \{x_2, y_2\}$$

Proof

$Var[x + y] = E[(x + y)^2] - \{E[x + y]\}^2$

$Var[x + y] = E[(x^2 + 2xy + y^2)] - \{E[x] + E[Y]\}^2$

$Var[x + y] = E[x^2] + E[y^2] + E[2xy] - \{E[x] + E[Y]\}^2$

$Var[x + y] = E[x^2] + E[y^2] + E[2xy] - E[x]^2 - E[y]^2 - 2E[x][y]$

$Var[x + y] = E[x^2] + E[y^2] + 2E[x][y] - E[x]^2 - E[y]^2 - 2E[x][y]$

$Var[x + y] = E[x^2] - E[x]^2 + E[y^2] - E[y]^2$

$Var[x + y] = Var[x] + E[y]$