4.2. Probability Distributions and Compound Lotteries

As some of the examples of the previous section suggest, we do not have to know the probability measure for all subsets of a sample space Ω. It is enough to know that probabilities for sets of the form $\{\omega\}$, where $\omega \in \Omega$.

4.2.1. Definition.

Let Ω be a sample space.

1. *A probability distribution is a function $\delta : \Omega \to [0,1]$ so that $\sum_{\omega \in \Omega} \delta(\omega) = 1$.*
2. *For each probability distribution δ and each subset $E \subseteq \Omega$ we define $\mu_\delta(E) = \sum_{\omega \in E} \delta(\omega)$.*
3. *If μ is a probability measure on Ω, then we define $\delta_\mu(\omega) = \mu(\{\omega\})$ for each $\omega \in \Omega$.*

4.2.2. Proposition.

Let Ω be a sample space.

1. *If $\delta : \Omega \to [0,1]$ is a probability distribution, then μ_δ is probability measure on Ω.*
2. *Conversely, if μ is a probability measure on Ω, then δ_μ is probability distribution.*
3. *We have $\mu_{(\delta_\mu)} = \mu$ and $\delta = \delta_{(\mu_\delta)}$.*

Proof. (1) By definition, $\mu_\delta(\varnothing) = \sum_{\omega \in \varnothing} \mu(\omega) = 0$ and $\mu_\delta(\Omega) = \sum_{\omega \in \Omega} \delta(\omega) = 1$. Moreover, if E and F are disjoint events, then $\mu_\delta(E \cup F) = \sum_{\omega \in E \cup F} \delta(\omega) = \sum_{\omega \in E} \delta(\omega) + \sum_{\omega \in F} \delta(\omega) = \mu_\delta(E) + \mu_\delta(F)$. Hence μ_δ is a probability measure.

(2) We compute:

$$\begin{aligned}\sum_{\omega \in \Omega} \delta_\mu(\omega) &= \sum_{\omega \in \Omega} \mu(\{\omega\}) \\ &= \mu\left(\bigcup_{\omega \in \Omega} \{\omega\}\right) \\ &= \mu(\Omega) \\ &= 1\end{aligned}$$

Hence δ_μ is a probability distribution.
The proof of (3) is left as an exercise.

4.2.3. Proposition.

If Ω is a sample space, if δ is a probability distribution on Ω and if f is a random variable, then the expected value of f with respect to the probability measure μ_δ is given by

$$E(f) = \sum_{\omega \in \Omega} f(\omega)\delta(\omega)$$

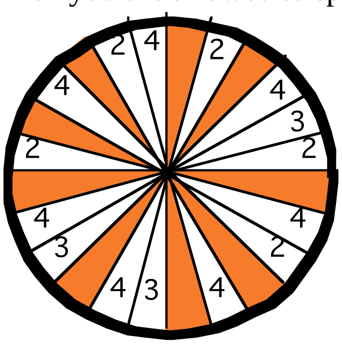
Proof. This follows immediately from the definitions:

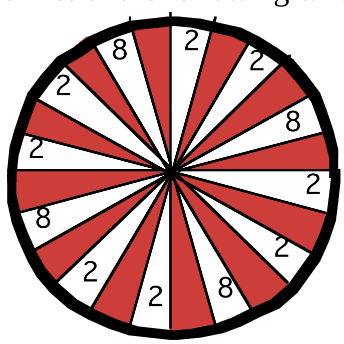
$$\begin{aligned} E(f) &= \sum_{\omega \in \Omega} f(\omega)\mu_\delta(\{\omega\}) \\ &= \sum_{\omega \in \Omega} f(\omega)\delta(\omega) \end{aligned}$$

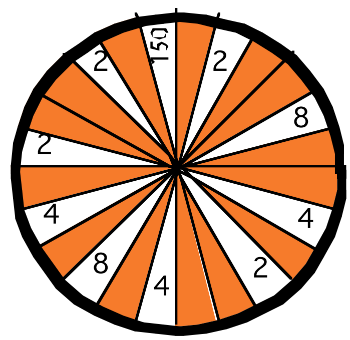
Before we continue, we consider the following example:

4.2.4. Example

At a country fair, one of the booths offers the following game: You buy a ticket for $2. Then you are allowed to spin the first of the following wheels:



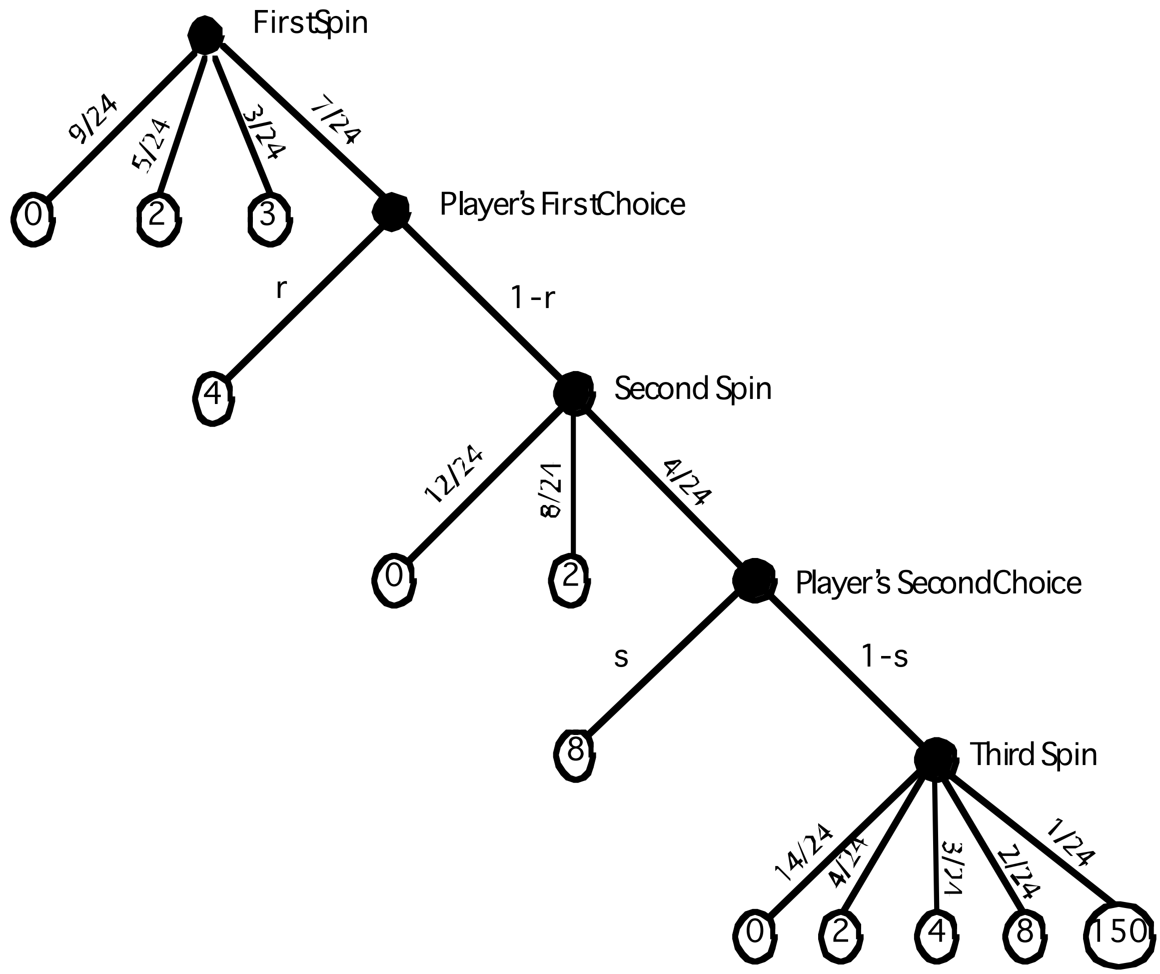




The wheel will stop in a certain field, and you will win the amount indicated on the wheel. If you win $4, you have the option to spin the second wheel, which has higher stakes. If you win $8 at the second spin, you can spin the third wheel, which tempts you with a win of $150. Then the game ends.

1. Find an optimal strategy to play this game.
2. If 1000 visitors of the fair play this game each day, how much money does the owner of the booth carry home at the end of a day? How does his income change if every 2nd player who has the option the spin the next wheel actually does so?

In order to answer this question, we create the game tree. This time, the game tree also has probabilities:



The nodes where the players have to make a choice do not have numerical values for the probabilities yet: We expect r % of the players to take the 4 dollars at the node representing the player's first choice, while (1-r) % of the players chose to continue the game and spin the next wheel.

At the third spin, the expected win of the player can be computed as

$$\frac{14}{24}\times 0+\frac{4}{24}\times 2+\frac{3}{24}\times 4+\frac{2}{24}\times 8+\frac{1}{24}\times 150=\frac{31}{4}$$

Only (1-s) % of the players will take this route, the remaining s % will cash in the win of $ 8. So the expected win at the node representing the Player's Second Choice have the value of

$$s\times 8+(1-s)\times\frac{186}{24}=\frac{1}{4}s+\frac{31}{4}$$

Hence the expected value at the Second Spin can be computed as

$$\frac{12}{24}\times 0+\frac{8}{24}\times 2+\frac{4}{24}\times\left(\frac{1}{4}s+\frac{31}{4}\right)=\frac{1}{24}s+\frac{47}{24}$$

Only (1-r) % of the players will make this choice, so the expected value at the node representing the Player's First choice is

$$4r+(1-r)\left(\frac{1}{24}s+\frac{47}{24}\right)=\frac{49}{24}r+\frac{1}{24}s-\frac{1}{24}rs+\frac{47}{24}$$

Finally, the total expected win can be computed as

$$\frac{9}{24}\times 0+\frac{5}{24}\times 2+\frac{3}{24}\times 3+\frac{7}{24}\times\left(\frac{49}{24}r+\frac{1}{24}s-\frac{1}{24}rs+\frac{47}{24}\right)$$
$$=\frac{343}{576}r+\frac{7}{576}s-\frac{7}{576}rs+\frac{785}{576}$$

If all players a cashing in their win at the first opportunity *(r=1)*, then the total expected win of the players would by given by \$47/24 = \$1.96, and the poor owner of the booth would keep in the average only 4 cent of each player. If 1000 players were playing the game, he would make only \$40, not nearly enough to feed his wife and children. If every second player decides to continue (*r = s = ½)*, then the expected win would be \$ 1.67, and owner of the booth would keep \$ 0.33 from each player. His total income for the day would be \$330.- If however, as expected, almost every player would continue (*r = s = 0.01*), then the expected win would be \$1.37, and he would keep \$630 for the day.

As we can see from the preceding example, the game tree now has probabilities. Some or the moves are decided by luck (the wheels), and the players determine the others. In order to define the expected value of a game, it is more convenient not to make distinctions between those to types.

4.2.5. Definition.

Let t be a tree. A compound lottery on t is a function λ that assigns a non-negative real number to each edge so that for each fixed node N of t the equation

$$\sum_{e\text{ starts at node }N}\lambda(e)=1$$

is valid.

In other words, for each node *N* of *t*, the function λ assigns a probability distribution to the set of all edges starting at *N*.

Next, we need to find the probability that a certain path starting at the root of a tree reaches a certain leaf. This probability can be defined as follows:

4.2.6. Definition.

Let t be a tree, and let λ be a compound lottery on t. For each leaf L of t let $P = e_1, e_2, \cdots, e_n$ be the unique path leading from the root R of t to the L. Then we define

$$\delta_\lambda(L) = \lambda(e_1)\lambda(e_2)\cdots\lambda(e_n)$$

We have to show that the function δ_λ is a probability distribution. In order to do this, we will use the recursive definition of trees. The next proposition serves also as a preparation for Zermelo's Algorithm.

4.2.7. Proposition.

Let t be a tree, and let λ be a compound lottery on t. Assume that $t = [t_1 \oplus \cdots \oplus t_n]$. For each index j let λ_j be the restriction of λ to the edges of the tree t_j. If L is a leaf of t, if k is the unique index so that L belongs to the tree t_k and if e is the unique edge starting at the root of R and ending at the root of t_k, then

$$\delta_\lambda(L) = \lambda(e)\delta_{\lambda_k}(L)$$

Proof. Let $P = e_1, e_2, \cdots, e_n$ be the unique path starting at the root R of *t* and ending at the leaf *L*. Then $e = e_k$ is the unique edge connecting *R* to the root of t_k, and $P_k = e_2, \cdots, e_n$ is the unique path in t_k from the root of t_k to *L*. Hence

$$\delta_{\lambda_k}(L) = \lambda(e_2)\lambda(e_3)\cdots\lambda(e_n)$$

and therefore

$$\begin{aligned}\delta_\lambda(L) &= \lambda(e_1)\lambda(e_2)\cdots\lambda(e_n)\\ &= \lambda(e)\delta_{\lambda_k}(L)\end{aligned}$$

4.2.8. Proposition.

Let t be a tree, and let λ be a compound lottery on t. Then δ_λ is a probability distribution on the set of all leaves of t, i.e.

$$\sum_{L \text{ is a leaf of } t} \delta_\lambda(L) = 1$$

Proof. We prove the theorem by induction on the number of edges. If there is no edge, then there is actually nothing to do. If there is exactly one edge *e*, then this edge connects the root of the tree to its only child, and this child is the only leaf *L* of the tree. Since λ is a probability distribution on the collection of all edges starting at the root *R*, and since there is only one such edge, it follows that $\lambda(e) = \sum_{e \text{ starts at R}} \lambda(e) = 1$, and hence $\sum_{L \text{ is a leaf of } t} \delta_\lambda(L) = \delta_\lambda(L) = \lambda(e) = 1$.

Assume that we have verified the proposition for all trees with *m* or fewer edges, and that *t* has *m+1* edges. Then $t = [t_1 \oplus \cdots \oplus t_n]$, and, since none of the children of *t* contains the root of *t*, each of the trees t_j has no more than *m* edges. It follows from the induction hypothesis that

$$\sum_{L \text{ is a leaf of } t_k} \delta_{\lambda_k}(L) = 1$$

Let e_i the unique edge connecting the root *R* of *t* with the root of t_i. Using the previous proposition and the fact that $e_1 + \cdots + e_n = \sum_{e \text{ starts at } R} \lambda(e) = 1$, we compute

$$\begin{aligned}
\sum_{L \text{ is a leaf of } t} \delta_\lambda(L) &= \sum_{i=1}^{n} \sum_{L \text{ is a leaf of } t_i} \delta_\lambda(L) \\
&= \sum_{i=1}^{n} \sum_{L \text{ is a leaf of } t_i} \lambda(e_i)\delta_{\lambda_i}(L) \\
&= \sum_{i=1}^{n} \lambda(e_i) \sum_{L \text{ is a leaf of } t_i} \delta_{\lambda_i}(L) \\
&= \sum_{i=1}^{n} \lambda(e_i) \times 1 \\
&= 1
\end{aligned}$$

Finally, we have to work out a formula for expected values. If Ω is the collection of all leaves of a tree with a compound lottery λ, and if $f : \Omega \to \Re$ is a random variable, then the expected value of *f* with respect to the probability distribution δ_λ was defined as

$$\begin{aligned}
E(f) &= \sum_{L \in \Omega} f(L)\delta_\lambda(L) \\
&= \sum_{L \text{ is a leaf of } t} f(L)\delta_\lambda(L)
\end{aligned}$$

The following theorem allows us to use a version of Zermelo's Algorithm to compute this value:

4.2.9. Theorem.

Let t be a tree with a compound lottery λ. *Let f be a random variable, defined on the leaves of t. Assume that*

1. *$t = [t_1 \oplus \cdots \oplus t_n]$ and that e_k is the unique edge connected the root of R with the root of t_k;*
2. *λ_k is the restriction of λ to t_k;*
3. *f_k is the restriction of f to the leaves of t_k;*
4. *$E(f_k)$ is the expected value of f_k, computed with respect to the probability distribution δ_{λ_k} on the leaves of t_k*

Then

$$E(f) = \lambda(e_1)E(f_1) + \cdots + \lambda(e_n)E(f_n)$$

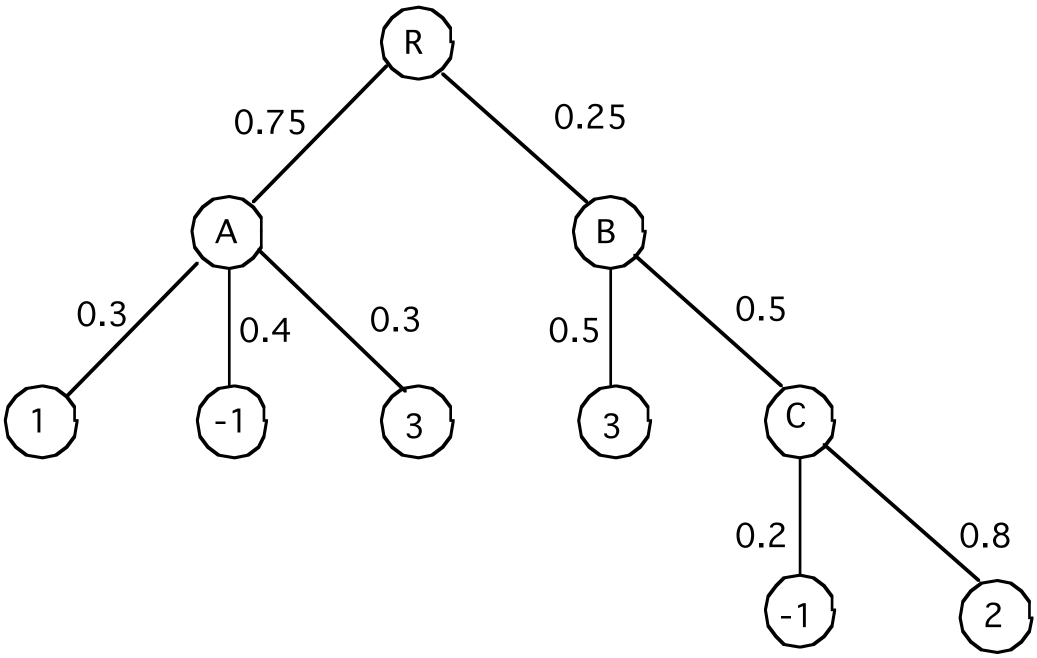
Proof. Using the definitions and previous propositions, we compute

$$\begin{aligned} E(f) &= \sum_{L \text{ is a leaf of } t} f(L)\delta_\lambda(L) \\ &= \sum_{k=1}^{n} \sum_{L \text{ is a leaf of } t_k} f(L)\delta_\lambda(L) \\ &= \sum_{k=1}^{n} \sum_{L \text{ is a leaf of } t_k} f_k(L)\lambda(e_k)\delta_{\lambda_k}(L) \\ &= \sum_{k=1}^{n} \lambda(e_k) \sum_{L \text{ is a leaf of } t_k} f_k(L)\delta_{\lambda_k}(L) \\ &= \sum_{k=1}^{n} \lambda(e_k)E(f_k) \end{aligned}$$

The last theorem allows us to use a version of Zermelo's Algorithm to compute expected values for compound lotteries, and we have use this method in preceding example.

4.2.10. Example

For the following tree, the values for the random variable are indicated at the leaves and the probabilities of the compound lottery are indicated at the edges. Find the expected value for the random variable, using Zermelo's Algorithm (i.e. the last theorem).



We compute the expected values for all subtrees, starting with the subtrees with root *A* and *C*, then moving to the subtree with root *B* and finally for the whole tree. Those values are:

1. For the tree starting at *A*: $0.3\times1+0.4\times(-1)+0.3\times3=0.8$
2. For the tree starting at *C*: $0.2\times(-1)+0.8\times2=1.4$
3. For the tree starting at *B*: $0.5\times3+0.5\times1.4=2.2$
4. For the whole tree: $0.75\times0.8+0.25\times2.2=1.15$