Using Discrete Probability Distributions

The slot machine has three independent windows and if all three windows light up in the same way, we win a prize.

Here are the probabilities of a particular image appearing in a particular window

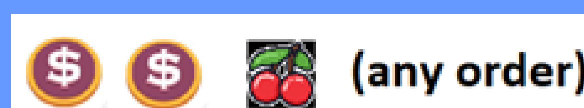
$	Cherry	Lemon	Other
0.1	0.2	0.2	0.5

$	Cherry	Lemon	Other
0.1	0.2	0.2	0.5

So let's work out some probabilities of winning.

Probability of 

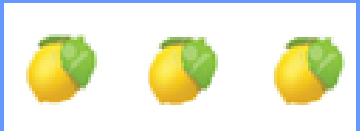
$$= 0.1 * 0.1 * 0.1$$
$$= 0.001$$

Probability of 

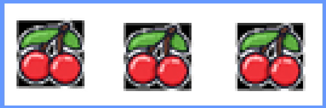
$$= (0.1 * 0.1 * 0.2) + (0.1 * 0.1 * 0.2) + (0.1 * 0.1 * 0.2)$$
$$= 0.006$$

$	Cherry	Lemon	Other
0.1	0.2	0.2	0.5

So let’s work out some probabilities of winning.

Probability of 

$$= 0.2 * 0.2 * 0.2$$
$$= 0.008$$

Probability of 

$$= 0.2 * 0.2 * 0.2$$
$$= 0.008$$

$	Cherry	Lemon	Other
0.1	0.2	0.2	0.5

Now what about the probability of losing?

$$= 1 - 0.001 - 0.006 - 0.008 - 0.008$$
$$= 0.977$$

We can compose a probability distribution for the slot machine.

Combination	**None**	**Lemons**	**Cherries**	**Dollars/ Cherry**	**Dollars**
Probability	0.977	0.008	0.008	0.006	0.001

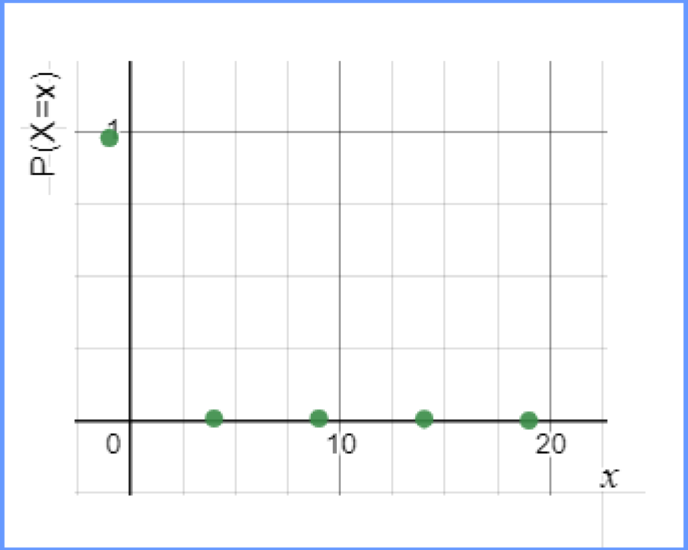
To make it more useful, let’s add a row showing how much we stand to win with each combination. (Remember it cost €1 to play)

Combination	None	Lemons	Cherries	Dollars/ Cherry	Dollars
Gain (x)	-€1	€4	€9	€14	€19
P(X=x)	0.977	0.008	0.008	0.006	0.001

This table gives us the **probability distribution** of the winnings, a set pf probabilities for every possible gain or loss.

P(X=x) means “The probability that the random variable **X** takes a particular value **x**”

Combination	None	Lemons	Cherries	Dollars/ Cherry	Dollars
Gain (x)	-€1	€4	€9	€14	€19
P(X=x)	0.977	0.008	0.008	0.006	0.001



We can also graph the probability distribution although this is of limited use in this case because of the massive difference between P(X=€-1) and the other values.

Once you have calculated a probability distribution, you can use this information to determine the expected outcome.

In the case of the slot machine, the probability distribution can be used to determine how much we can expect to win or lose long-term.

In other words we can find the **expectation**

The expectation is like the mean, except for probability distributions, like the mean, it is sometimes given the letter μ, but more usually, E.

Here is the equation for working out E.

$$\mathbf{E}= \sum \boldsymbol{x}\boldsymbol{P}(\boldsymbol{X} = \boldsymbol{x})$$

Let’s use it to calculate the expectation of the slot machine.

Combination	**None**	**Lemons**	**Cherries**	**Dollars/ Cherry**	**Dollars**
Gain (x)	-€1	€4	€9	€14	€19
P(X=x)	0.977	0.008	0.008	0.006	0.001

$$\mathbf{E}= (-\mathbf{1} \times \mathbf{0.977}) + (\mathbf{4} \times \mathbf{0.008}) + (\mathbf{9} \times \mathbf{0.008}) + (\mathbf{14} \times \mathbf{0.006}) + (\mathbf{19} \times \mathbf{0.001})$$

= -0.77

Here it is done on Excel.

x	-1	4	9	14	19	
P(X=x)	1	0.977	0.008	0.006	0.001	
	-1	0.032	0.072	0.084	0.019	-0.77

Here it is done in Octave/Matlab.

```
x=[-1,4,9,14,19];
P=[0.977,0.008,0.008,0.006,0.001];
E=sum(x.*P)
E = -0.7700
```

However you do it, the expectation is that you will lose €0.77 for each game.

Variance

Like the mean, the expectation doesn’t give the full story as the amount you stand to win on each game could vary a lot.

The expectation gives the average value of a variable but it doesn’t tell us how the values are spread out. To do this we use the **variance**.

Variance is the square of the average distance from the mean so let’s proceed along those lines.

$$Var(X) = \sum (x - E)^2 P(X = x)$$

3. Sum up the values

1. Calculate distance from the mean and square it.

2. Multiply each one by the probability x

Here it is in Excel.

x	-1	4	9	14	19	
P(X=x)	0.977	0.008	0.008	0.006	0.001	
xP(X=x)	-0.977	0.032	0.072	0.084	0.019	-0.77
	0.051683	0.182	0.764	1.3089	0.391	2.6971

E

Var(X)

Here it is in Octave/Matlab.

```
x=[-1,4,9,14,19];
P=[0.977,0.008,0.008,0.006,0.001];
E=sum(x.*P);
Var=sum(((x-E).^2).*P)
```

Standard Deviation

Probabilities also have a standard deviation

$$\boldsymbol{\sigma} = \sqrt{\boldsymbol{Var(X)}}$$

So the standard deviation of the slot machine winnings is $\sqrt{2.6971} = 1.642$

This means that on average our winnings per game will be €1.642 away from the expectation of €-0.77

Example 1

Here is the probability distribution of a random variable X

x	1	2	3	4	5
P(X=x)	0.1	0.25	0.356	0.2	0.1

Calculate

1. The expectation E(X)
2. The variance Var(X)

x	1	2	3	4	5
P(X=x)	0.1	0.25	0.356	0.2	0.1

To calculate the expectation E(X) we use

$$\boldsymbol{E(X) = \sum x\, P(X = x)}$$

E=(1x0.1)+(2x0.25)+(3x0.35)+(4x0.2)+(5x0.1)

=2.95

x	1	2	3	4	5
P(X=x)	0.1	0.25	0.356	0.2	0.1

To calculate the variance Var(X) we use

$$\mathbf{Var}(\boldsymbol{X}) = \sum(\boldsymbol{x} - \boldsymbol{E})^{\mathbf{2}}\, \boldsymbol{P}(\boldsymbol{X} = \boldsymbol{x})$$

$\text{Var}(X)=(1-2.95)^2\text{x}0.1+(2-2.95)^2\text{x}0.25+(3-2.95)^2\text{x}0.35+$
$(4-2.95)^2\text{x}0.2+ (5-2.95)^2\text{x}0.1$

$=1.2475$

Supposing the parameters change.

Here is the new probability distribution.

Combination	None	Lemons	Cherries	Dollars/ Cherry	Dollars
Gain (y)	-€2	€23	€48	€73	€98
P(Y=y)	0.977	0.008	0.008	0.006	0.001

Combination	None	Lemons	Cherries	Dollars/ Cherry	Dollars
Gain (y)	-€2	€23	€48	€73	€98
P(Y=y)	0.977	0.008	0.008	0.006	0.001

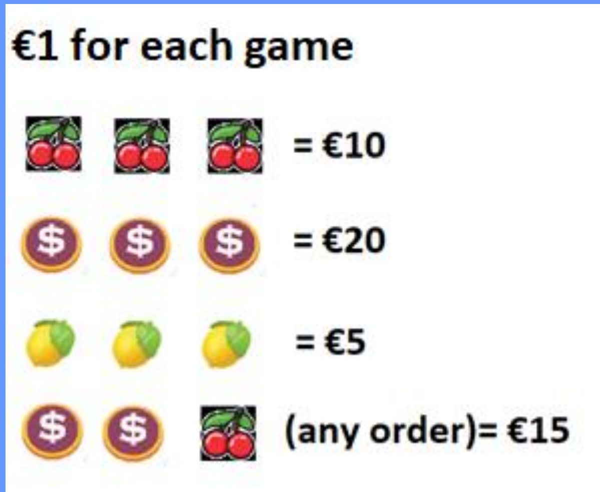
Let's work out the expectation and variation for the new conditions.

```
x=[-2,23,48,73,98];
P=[0.977,0.008,0.008,0.006,0.001];
E=sum(x.*P)
```

E = -0.8500

```
x=[-2,23,48,73,98];
P=[0.977,0.008,0.008,0.006,0.001];
E=sum(x.*P);
Var=sum(((x-E).^2).*P)
```

Var=67.428



E = -0.77

Var=2.6971

E = -0.8500

Var=67.428

The expectation is slightly lower so that in the long term we can expect to lose €0.85 each game.

The variance is much larger meaning that there is less certainty about the expectation.

So far we have looked at the probability distribution of various outcomes in a random betting process.

We calculated the mean of the probability distribution so we could find the expectation of winning in the long term, i.e. after many trials.

We also calculated the variance of the probability distribution so we could find the spread of expectations of winning in the long term.

Now we are going to examine another type of probability distribution.

Geometric Probability Distribution

Consider the case of a novice snowboarder. The probability of him making a successful run down the slope is 0.2

He has decided to keep trying until he has a successful run and then he will stop.

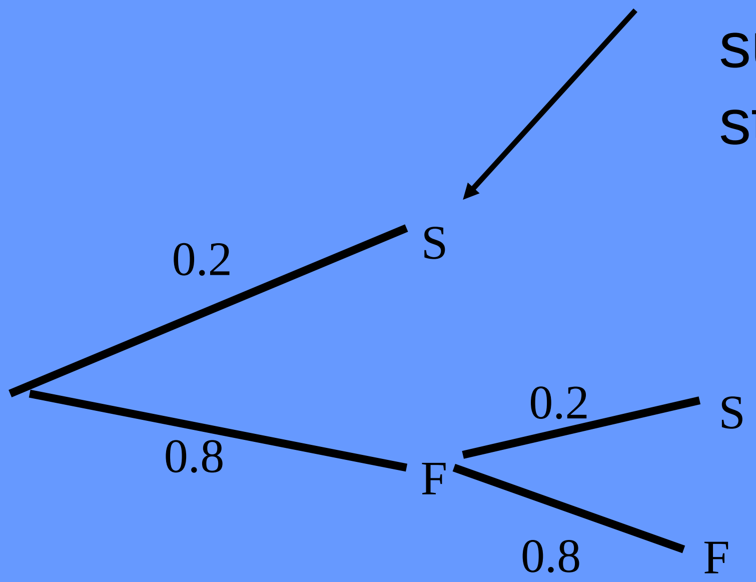
For the sake of simplicity, we will assume that he doesn’t improve with practice so the probability of success with each run remains constant at 0.2. (i.e. the trials are independent)

If X=number of trials, what is the probability of him making a successful run after x=2?

We will use a tree diagram

S=success, F-fail

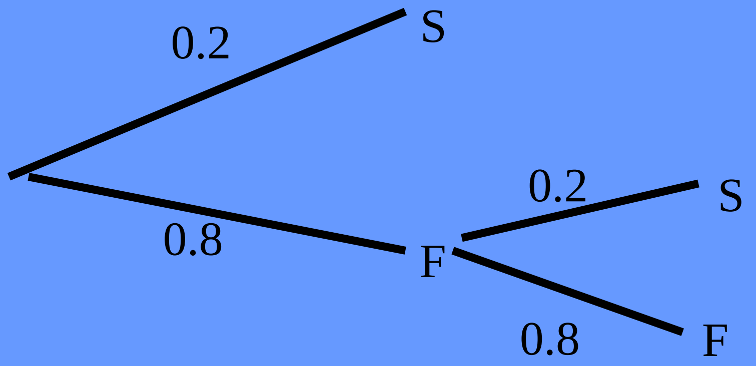
As soon as he succeeds, he stops



Trial 1 $P(X=1)= 0.2$

Trial 2 $P(X=2)= 0.2 \times 0.8 = 0.16$

S=success, F-fail



Trial 1 $P(X=1) = 0.2$

Trial 2 $P(X=2) = 0.2 \times 0.8 = 0.16$

$$P(X \leq 2) = 0.2 + 0.16 = 0.36$$

We can add these probabilities because they're independent.

Supposing we wanted to find the probability that our snowboarder would need say 10 or 20 attempts for a successful descent.

X	P(X success)
1	0.2
2	0.8×0.2=0.16
3	0.8×0.8×0.20=128
4	0.8×0.8×0.8×0.2=0.1024

$$P(r) = 0.8^{r-1} \times 0.2$$

$$P(X = r) = 0.8^{r-1} \times 0.2$$

$$P(10) = 0.8^{9} \times 0.2$$

$$= 0.8^{9} \times 0.2 = 0.027$$

P(X=r) means “The probability that X equals the value r” where r is the number of trials needed to get the first success.

We can generalise the formula further as

$$P(X=r)= pq^{r-1}$$

p=probability of success and

q=1-p=probability of failure.

This **geometric progression** covers situations where

- You run a series of independent trials.
- There can be a success or a failure for each trial
- We are interested in finding how many trials are needed to get the first successful outcome.

Example 2

The probability that a bit transmitted through a digital transmission channel is received in error is 0.1. Assume the transmissions are independent events, and let the random variable X denote the number of bits transmitted *until* the first error.

Calculate the probability that the first four bits are transmitted correctly and the fifth bit is in error.

$$P(X=r) = pq^{r-1}$$

$$P(X=5) = (0.1)(0.9^4) = 0.066$$

Note that P, the probability of “success” in this case refers to the probability of the bit being transmitted in error.

“Success” in this context means, “the event that we are looking for.”

Going back to the case of the snowboarder

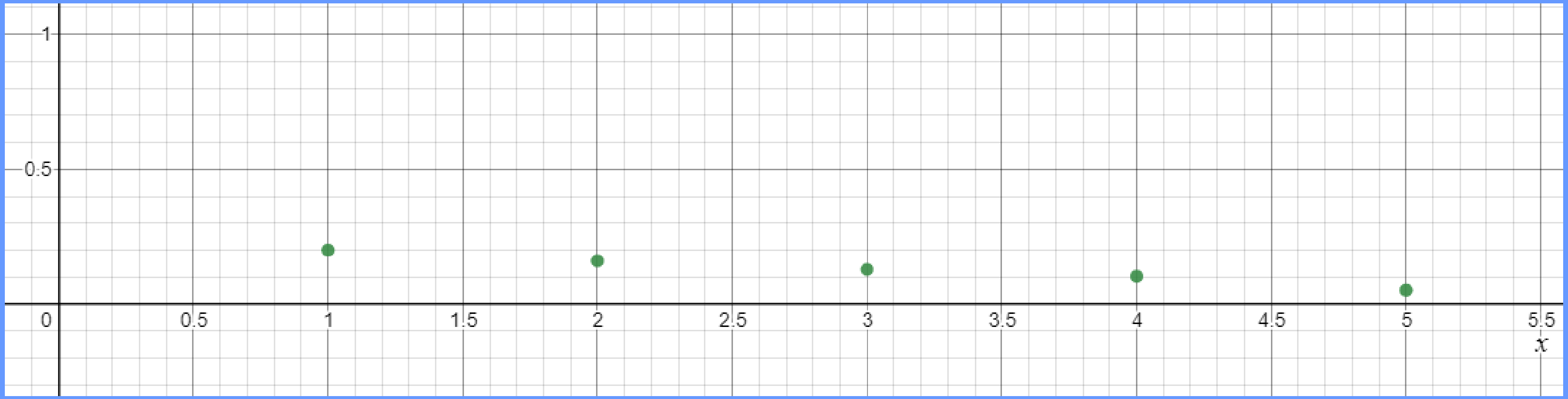
$$P(X = r) = 0.8^{r-1} \times 0.2$$

Let's work out the probability of success for a succession of trials

P(X=r)	**0.8^{r-1}*0.2**	
P(X=1)	0.8^{0}*0.2	0.2
P(X=2)	0.8^{1}*0.2	0.16
P(X=3)	0.8^{2}*0.2	0.128
P(X=4)	0.8^{3}*0.2	0.1024
P(X=5)	0.8^{4}*0.2	0.0512

Plotting the data shows the general shape of the geometric progression.

P(X=r) is at it’s highest when r=1 and it gets lower as r increases.



So it is most likely that only one trial will be required for a successful outcome.

The geometric distribution also works for inequalities

e.g. to find the probability that P(X>3). This means that the snowboarder failed three times so the probability is the probability of failure multiplied together three times = q^3

P(X=r)	**$0.8^{r-1}*0.2$**	
P(X=1)	$0.8^0*0.2$	0.2
P(X=2)	$0.8^1*0.2$	0.16
P(X=3)	$0.8^2*0.2$	0.128
P(X=4)	$0.8^3*0.2$	0.1024
P(X=5)	$0.8^4*0.2$	0.0512

$$P(X > 3) = 0.8^3$$

$$=0.512$$

Similarly $P(X \leq 3)$

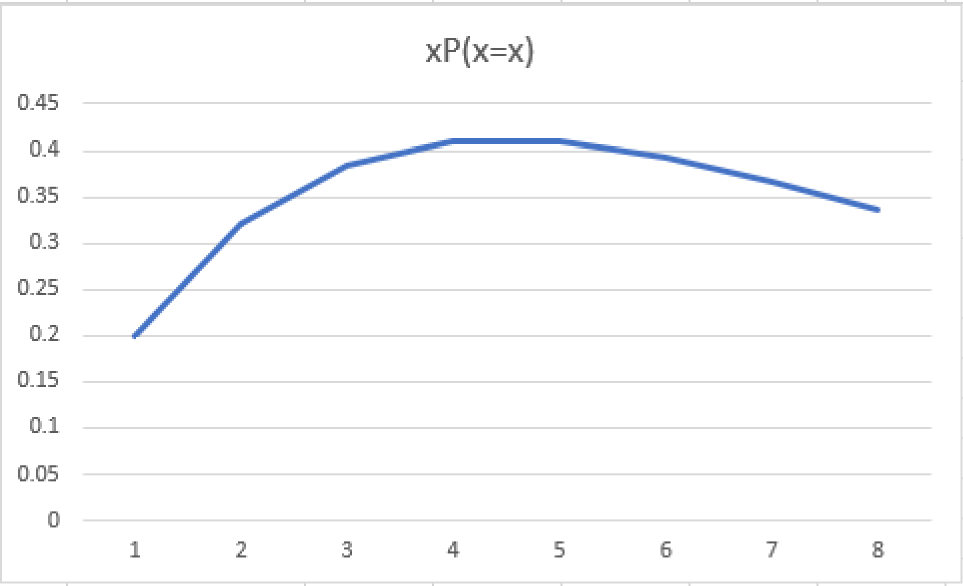
$$= 1 - 0.8^3$$

=0.488

x	P(X=x)		xP(x=x)	xP(X≤x)
1	0.2	0.2	0.2	0.2
2	0.8x0.2=0.16	0.16	0.32	0.52
3	0.8^2x0.2=0.128	0.128	0.384	0.904
4	0.8^3x0.2=0.1024	0.1024	0.41	1.3136
5	0.8^4x0.2=0.08192	0.08192	0.41	1.7232
6	0.8^5x0.2=0.065536	0.065536	0.393	2.11642
7	0.8^6x0.2=0.0524288	0.0524288	0.367	2.48342
8	0.8^7x0.2=0.04194304	0.04194304	0.336	2.81896

This is because the probability of the snowboarder failing 3 or less times is the complement of him or her failing more than 3 times

x	xP(x=x)
1	0.2
2	0.32
3	0.384
4	0.41
5	0.41
6	0.393
7	0.367
8	0.336



The values of the expectation xP(X=x) start off small and then get larger until x=5

x	xP(X≤x)
1	0.2
2	0.52
3	0.904
4	1.3136
5	1.7232
6	2.11642
7	2.48342
8	2.81896
9	3.12095
10	3.38939
11	3.62561
12	3.83177
13	4.01044
14	4.16437

The values of xP(X≤x), the expectation, gets closer and closer to a particular value, 5 in this case.

This makes intuitive sense. The probability of success is 0.2 so we can expect the snowboarder to make $\frac{1}{0.2} = 5$ attempts before he is successful.

$$E(X) = \frac{1}{p}$$

A similar process helps us to derive an expression for the variance of a geometric distribution

$$Var(X) = \frac{q}{p^2}$$

Example 3

We will finish by looking at one more example to help reinforce the meaning of the term Expectation as the mean of a probability distribution

One thousand tickets are sold at €1 each for a laptop valued at €350

What is the expected value of the gain if a person purchases one ticket?

One thousand tickets are sold at €1 each for a laptop valued at €350

What is the expected value of the gain if a person purchases one ticket?

	Win	Lose
Gain X	€349	-€1
Probability P(X)	1/1000	999/1000

	Win	Lose
Gain X	€349	-€1
Probability P(X)	1/1000	999/1000

Recall $\mathbf{E} = \sum \boldsymbol{x} \boldsymbol{P}(\boldsymbol{X} = \boldsymbol{x})$

$$\mathbf{E(X)} = \mathbf{349} * \frac{\mathbf{1}}{\mathbf{1000}} + (-\mathbf{1}) * \frac{\mathbf{999}}{\mathbf{1000}} = -€\mathbf{0.65}$$

Note that the expectation is -€0.65 does not mean that a person loses €0.65 since the person can only win a laptop worth €350 or lose €1 on the ticket.

What this expectation means is that the average of the losses is €0.65 for each of the 1000 ticket holders.

Another perspective is that if a person purchased one ticket a week over a long time, the average loss would be €0.65 since, in theory the person would win the laptop once for each 1000 tickets purchased.