Expectation

Let $X : S \to T$ be a discrete random variable (so T is countable). Then the expected value (or average) of X is written as $E[X]$ and is given by

$$E[X] = \sum_{t \in T} t \cdot P(X = t)$$

provided that the sum converges absolutely. In this case we say that X has "finite expectation".

- If the sum diverges to $\pm\infty$ we say the random variable has infinite expectation.
- If the sum diverges, but not to infinity, we say the expected value is undefined.

Properties of Expected Value

Suppose that X and Y are discrete random variables, both with finite expected value and both defined on the same sample space S. If a and b are real numbers then

$$E[aX] = aE[X];$$

$$E[X + Y] = E[X] + E[Y]; \quad \text{and}$$

$$E[aX + bY] = aE[X] + bE[Y].$$

If $X \geq 0$ then $E[X] \geq 0$.

Variance

Let $X : S \to T$ be a discrete random variable with finite expected value. Then the variance of the random variable is written as $Var[X]$ and is defined as

$$Var[X] = E[(X - E[X])^2] = \sum_{t \in T} (t - E[X])^2) P(X = t)$$

- The standard deviation of X is written as $SD[X]$ and is defined as

$$SD[X] = \sqrt{Var[X]}$$

Interpretation of Variance

- If X has a high probability of being far away from $E[X]$ the variance will tend to be large, while if X is very near $E[X]$ with high probability the variance will tend to be small.
- If we were to associate units with the random variable X (say *meters*), then the units of $Var[X]$ would be $meters^2$ and the units of $SD[X]$ would be *meters*.
- Informally we will view the standard deviation as a typical distance from average.
- It is possible that $Var[X]$ and $SD[X]$ could be infinite even if $E[X]$ is finite – meaning that the random variable has a clear average, but is so spread out that any finite number underestimates the typical distance of the random variable from its average.

Standardising Random Variables

- A standardized random variable X is one for which

$$E[X] = 0 \quad \text{and} \quad Var[X] = 1.$$

- Let X be a discrete random variable with finite expected value and finite, non-zero variance. Then $Z = \frac{X - E[X]}{SD[X]}$ is a standardized random variable.

Sampling from a given distribution

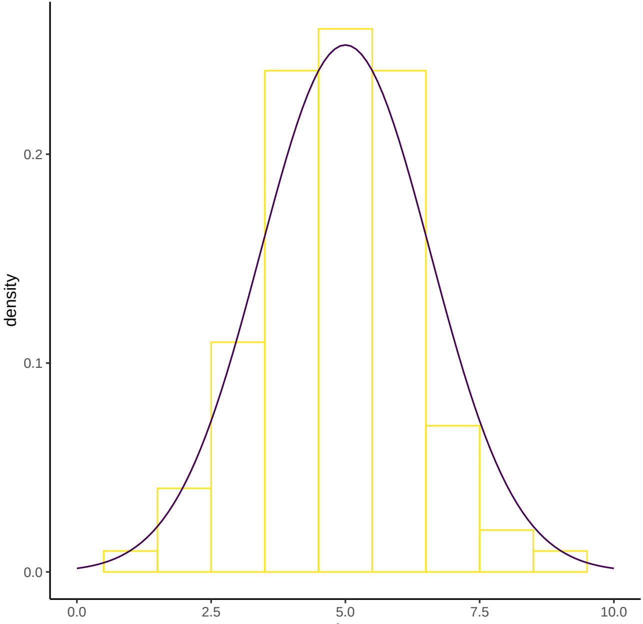
- we can use the `sample` function.
- takes a sample of the specified size (specified by `size`) from the elements of x using either with or without replacement (specified by `replace`). The optional `prob` argument can be used to give a vector of weights for obtaining the elements of the vector being sampled.

```
> x = c(1,2,3,4,5,6)
> probx= c(1/6,1/6,1/6,1/6,1/6,1/6)
> Rolls=sample(x, size=1800,  replace=T, prob=probx)
```

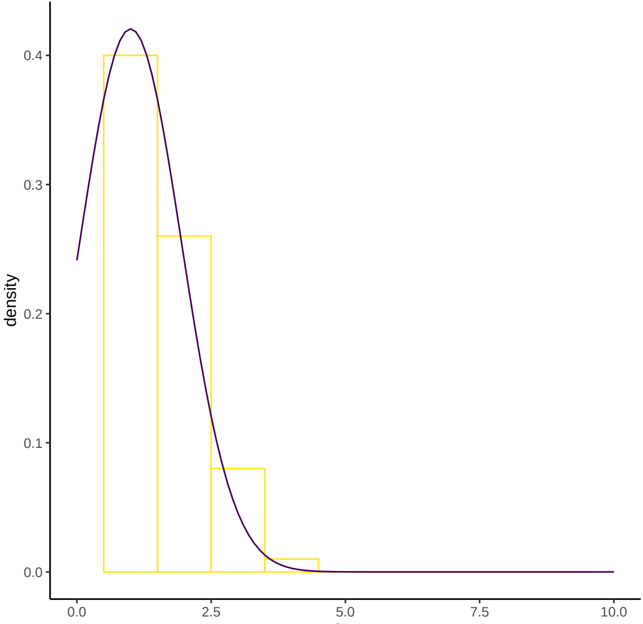
Uniform(1,2,3,4,5,6)

```
> mean(Rolls)
[1] 3.511111
> var(Rolls)
[1] 2.80588
```

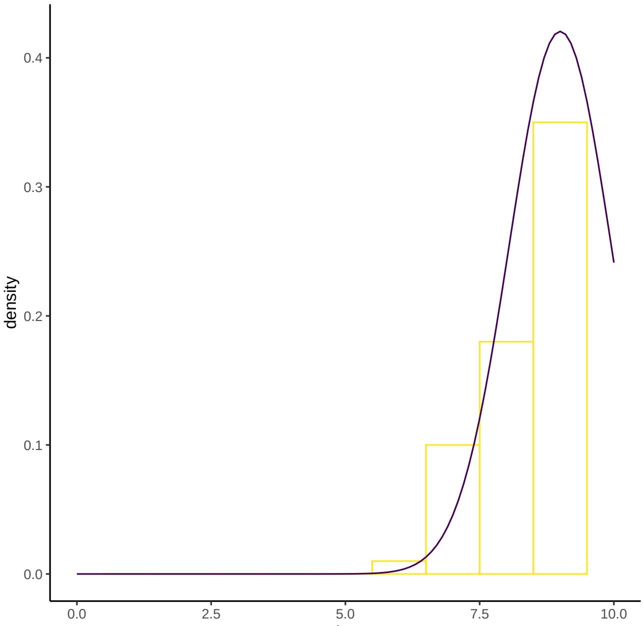
Binomial $(10, 0.5)$



Binomial $(10, 0.1)$



Binomial $(10, 0.9)$



Sums of Rolls