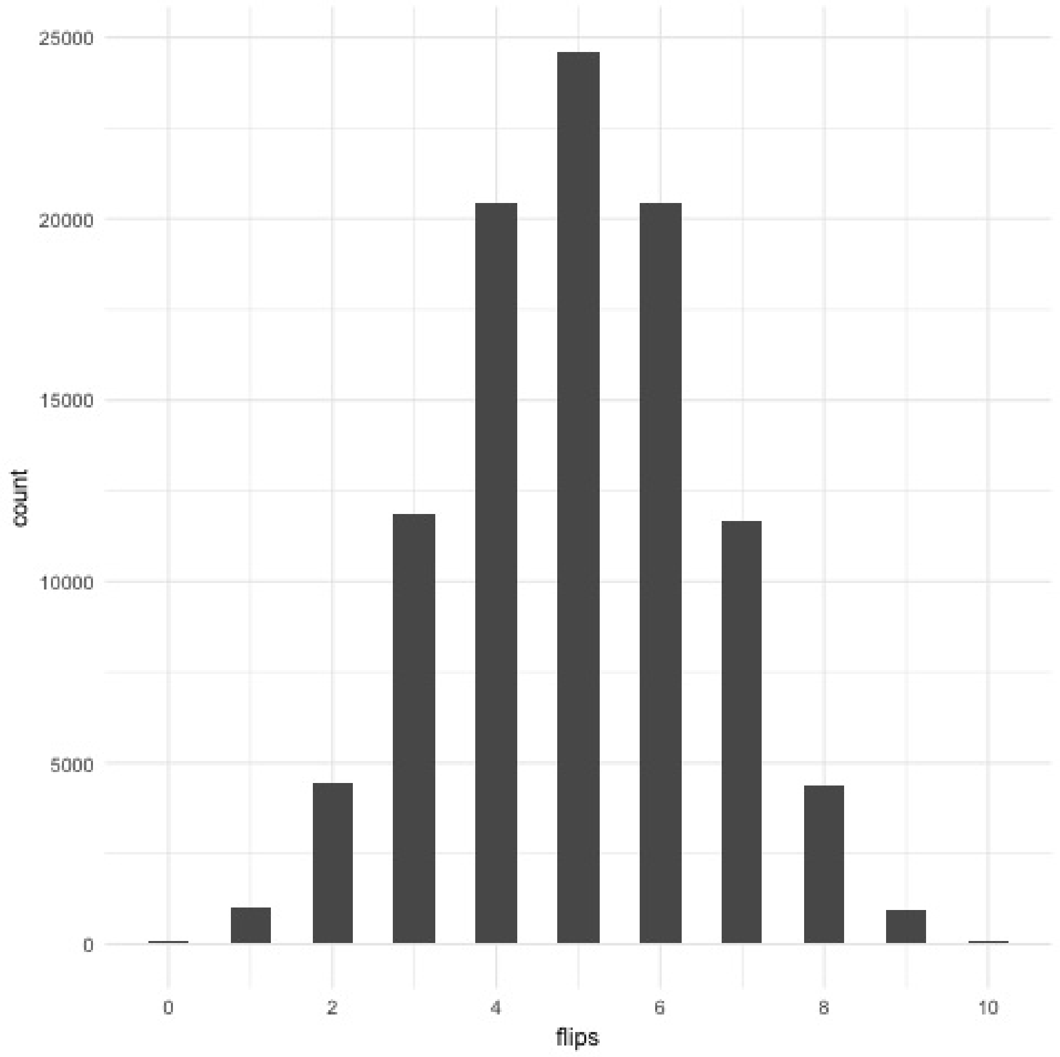
FOUNDATIONS OF PROBABILITY IN R

The normal distribution

David Robinson
Data Scientist, Stack Overflow

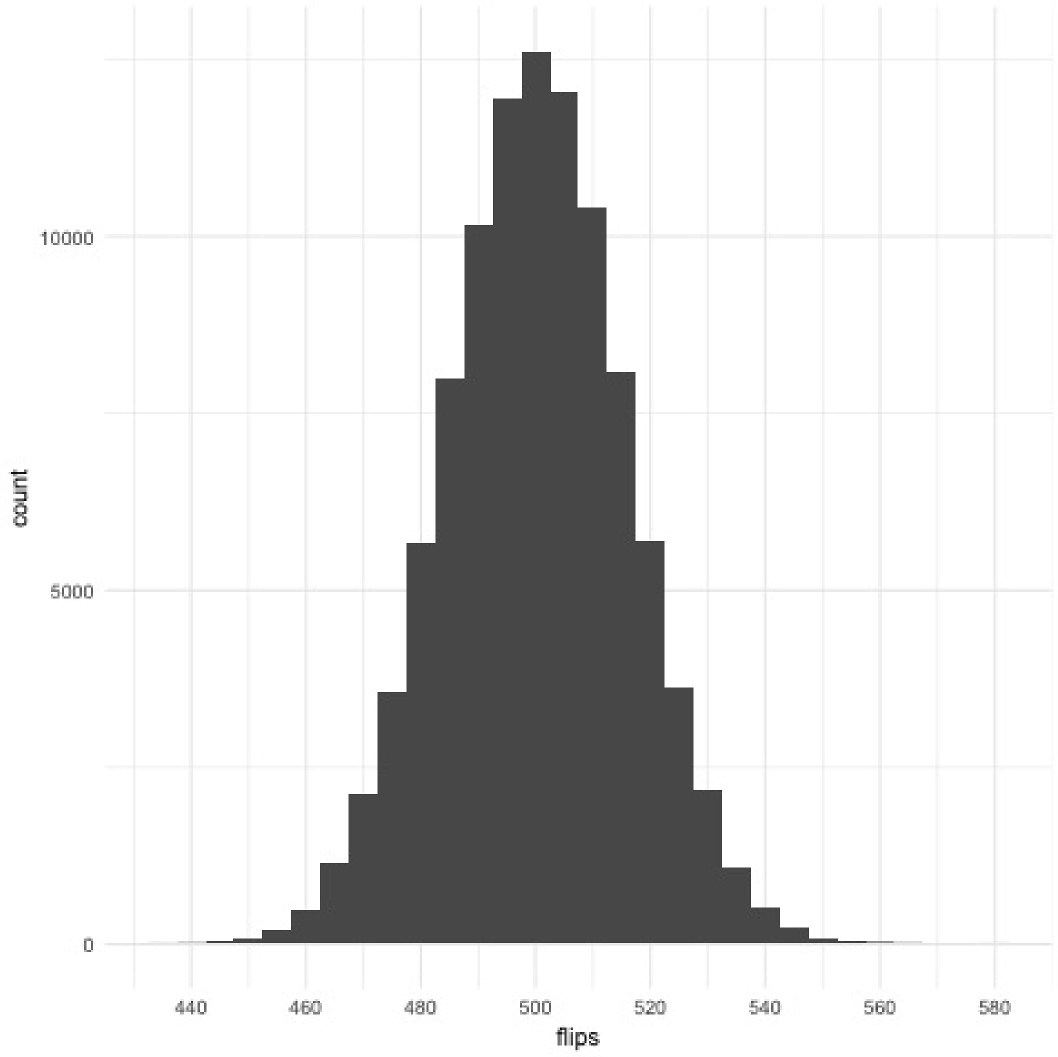
Flipping 10 coins

```
flips <- rbinom(100000, 10, .5)
```

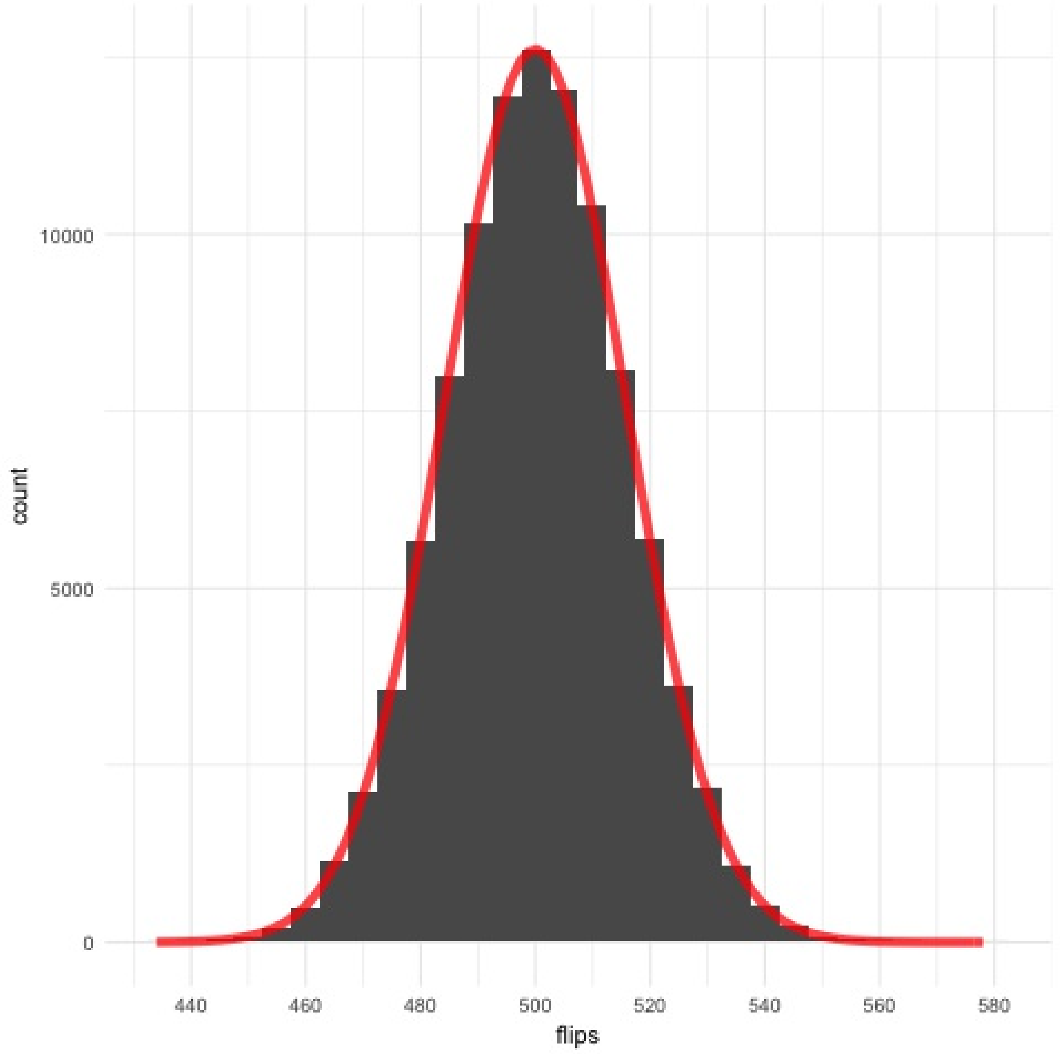


Flipping 1000 coins

```
flips <- rbinom(100000, 1000, .5)
```



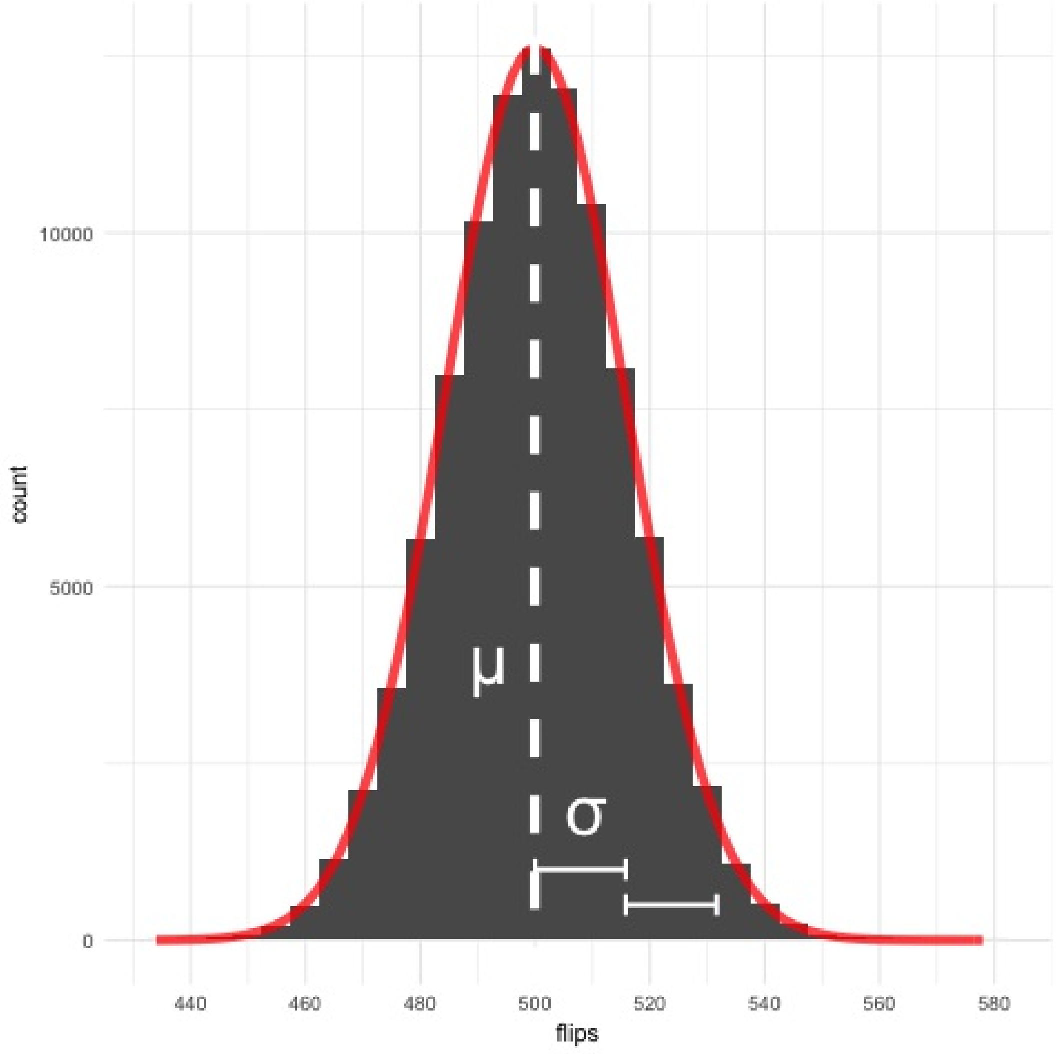
Flipping 1000 coins



Normal distribution has mean and standard deviation

$$X \sim \mathrm{Normal}(\mu, \sigma)$$

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



Normal approximation to the binomial

```
binomial <- rbinom(100000, 1000, .5)
```

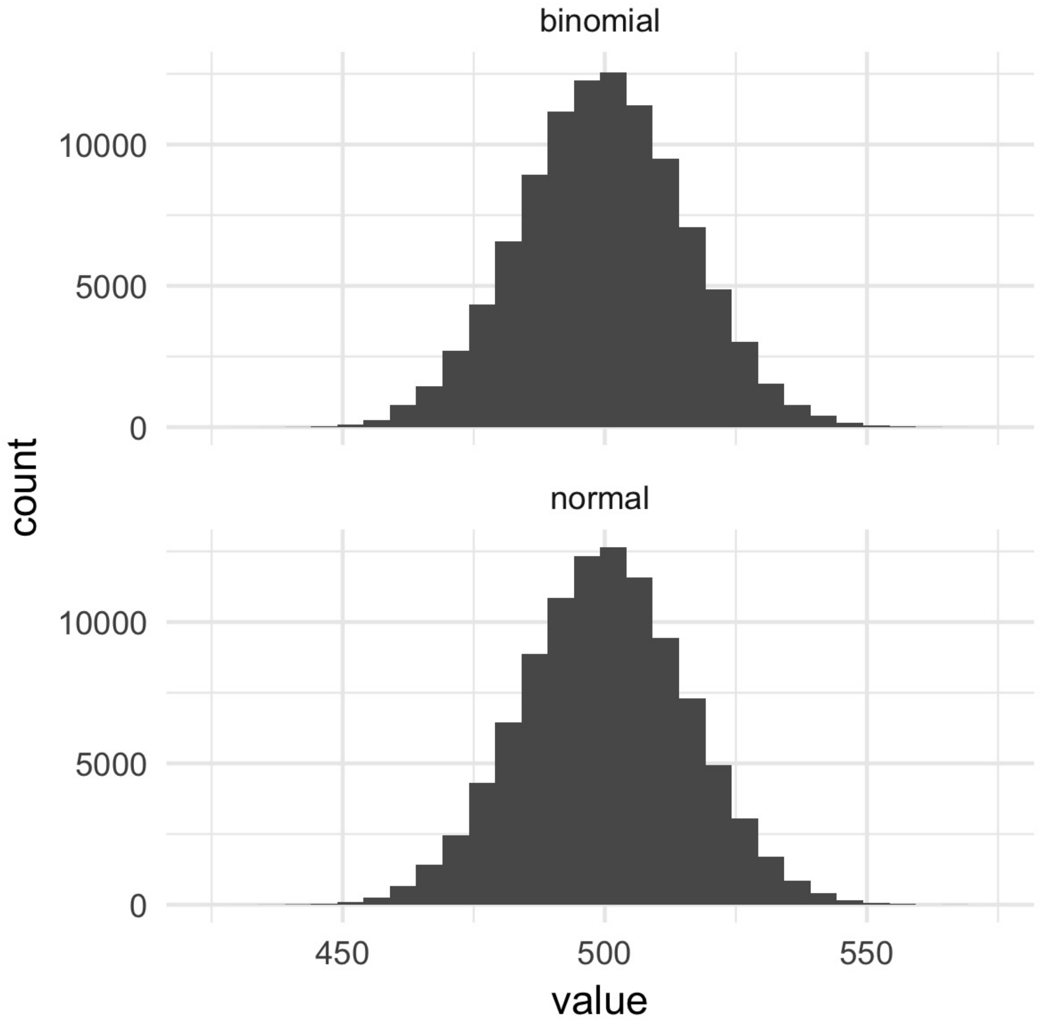
$$\mu = \text{size} \cdot p$$
$$\sigma = \sqrt{\text{size} \cdot p \cdot (1 - p)}$$

```
expected_value <- 1000 * .5
variance <- 1000 * .5 * (1 - .5)
stdev <- sqrt(variance)
```

```
normal <- rnorm(100000, expected_value, stdev)
```

Comparing histograms

```
compare_histograms(binomial, normal)
```



FOUNDATIONS OF PROBABILITY IN R

Let's practice!

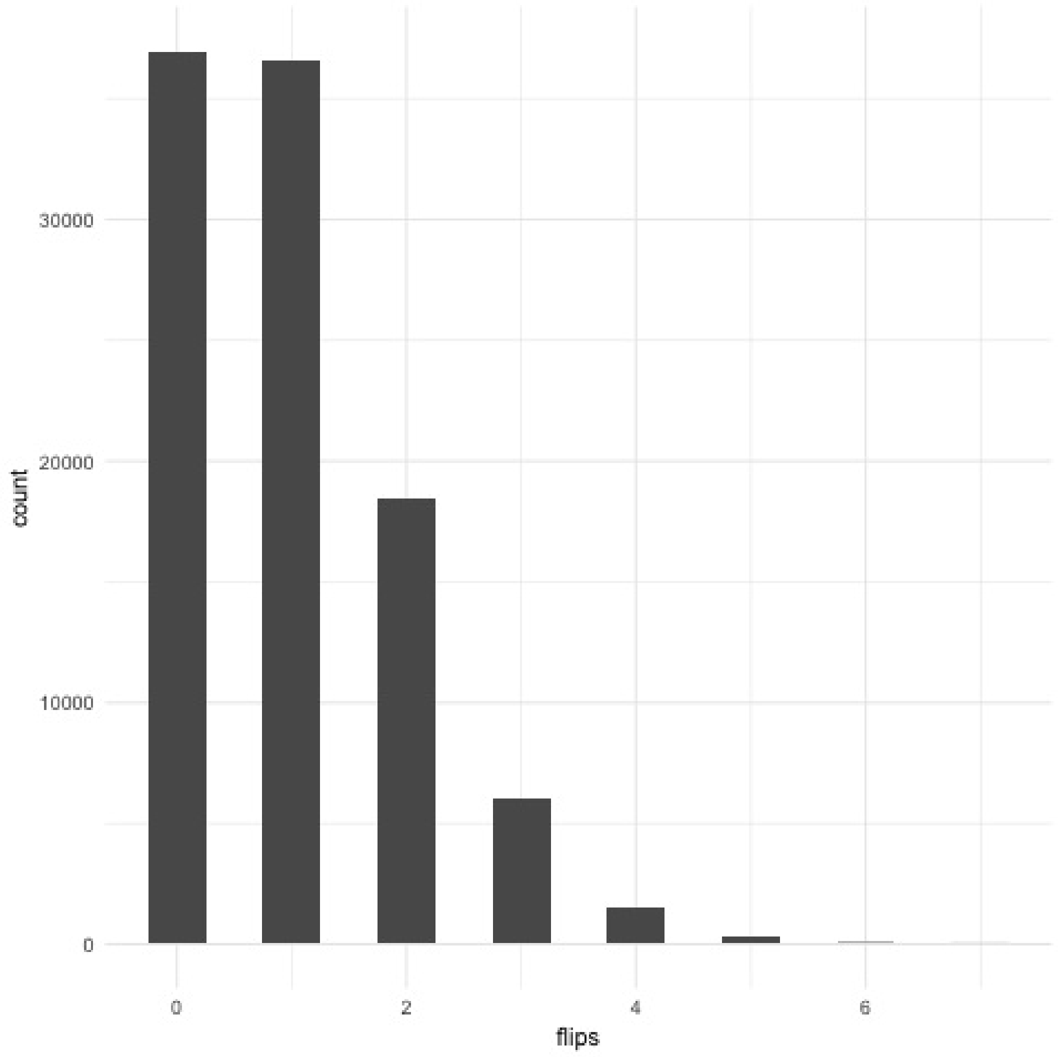
FOUNDATIONS OF PROBABILITY IN R

The Poisson distribution

David Robinson
Data Scientist, Stack Overflow

Flipping many coins, each with low probability

```
binomial <- rbinom(100000, 1000, 1 / 1000)
```



Properties of the Poisson distribution

```
binomial <- rbinom(100000, 1000,
                   1 / 1000)
```

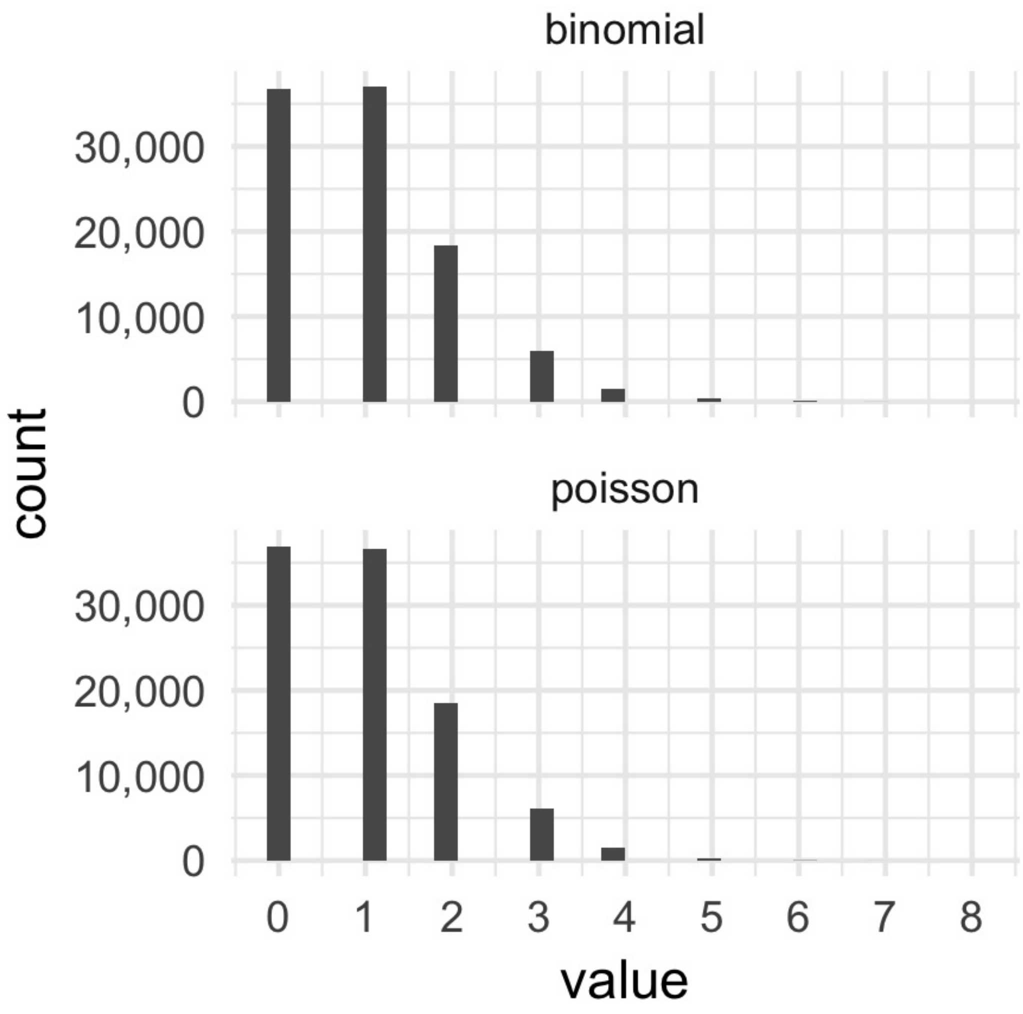
```
poisson <- rpois(100000, 1)
```

```
compare_histograms(binomial,
                   poisson)
```

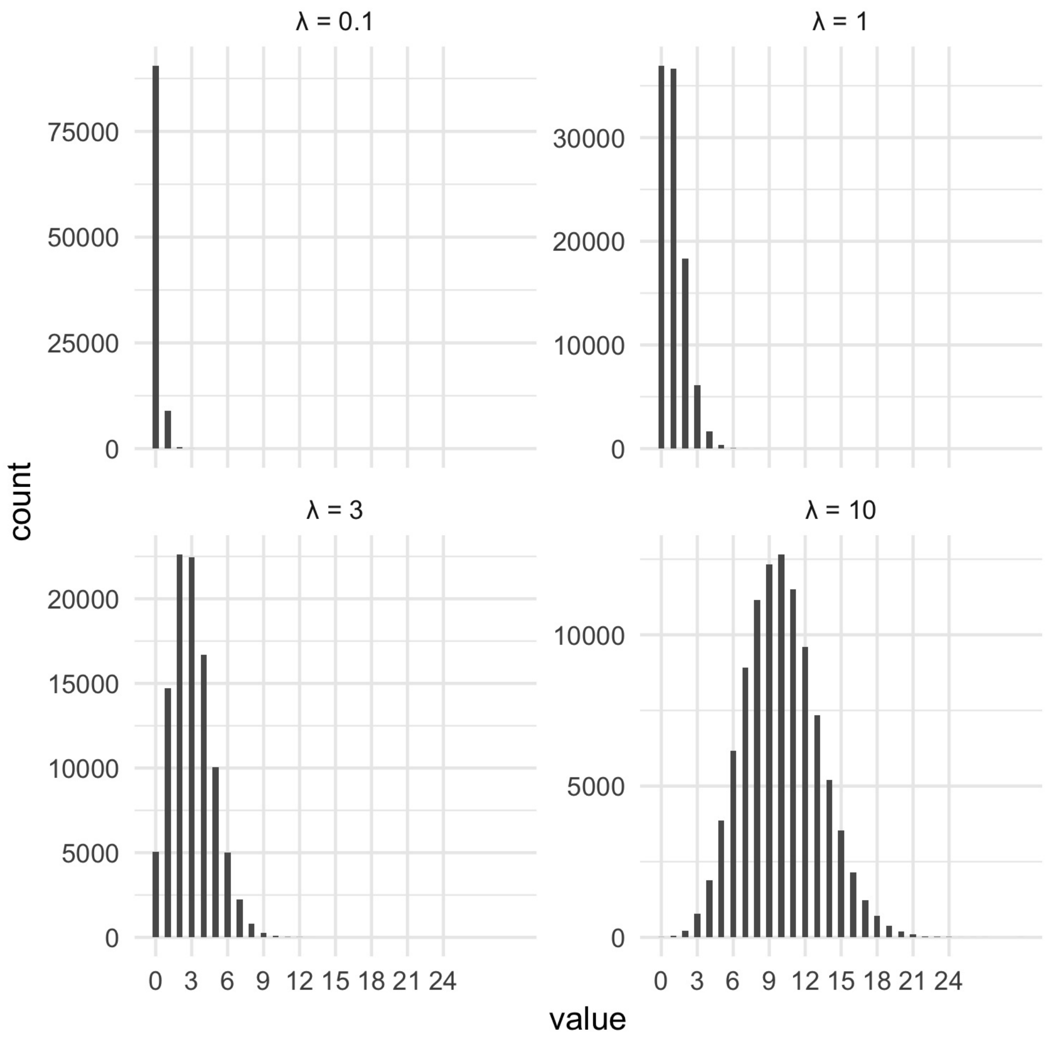
$$X \sim \text{Poisson}(\lambda)$$

$$E[X] = \lambda$$

$$\text{Var}(X) = \lambda$$



Poisson distribution



FOUNDATIONS OF PROBABILITY IN R

Let's practice!

FOUNDATIONS OF PROBABILITY IN R

The geometric distribution

David Robinson
Data Scientist, Stack Overflow

Simulating waiting for heads

```
flips <- rbinom(100, 1, .1)
flips
#  [1] 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0
# [16] 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0
```

```
which(flips == 1)
# [1] 8 27 44 55 82 89
```

```
which(flips == 1)[1]
# [1] 8
```

Replicating simulations

```
which(rbinom(100, 1, .1) == 1)[1]
# [1] 28
```

```
which(rbinom(100, 1, .1) == 1)[1]
# [1] 4
```

```
which(rbinom(100, 1, .1) == 1)[1]
# [1] 11
```

```
replicate(10, which(rbinom(100, 1, .1) == 1)[1])
# [1] 22 12  6  7 35  2  4 44  4  2
```

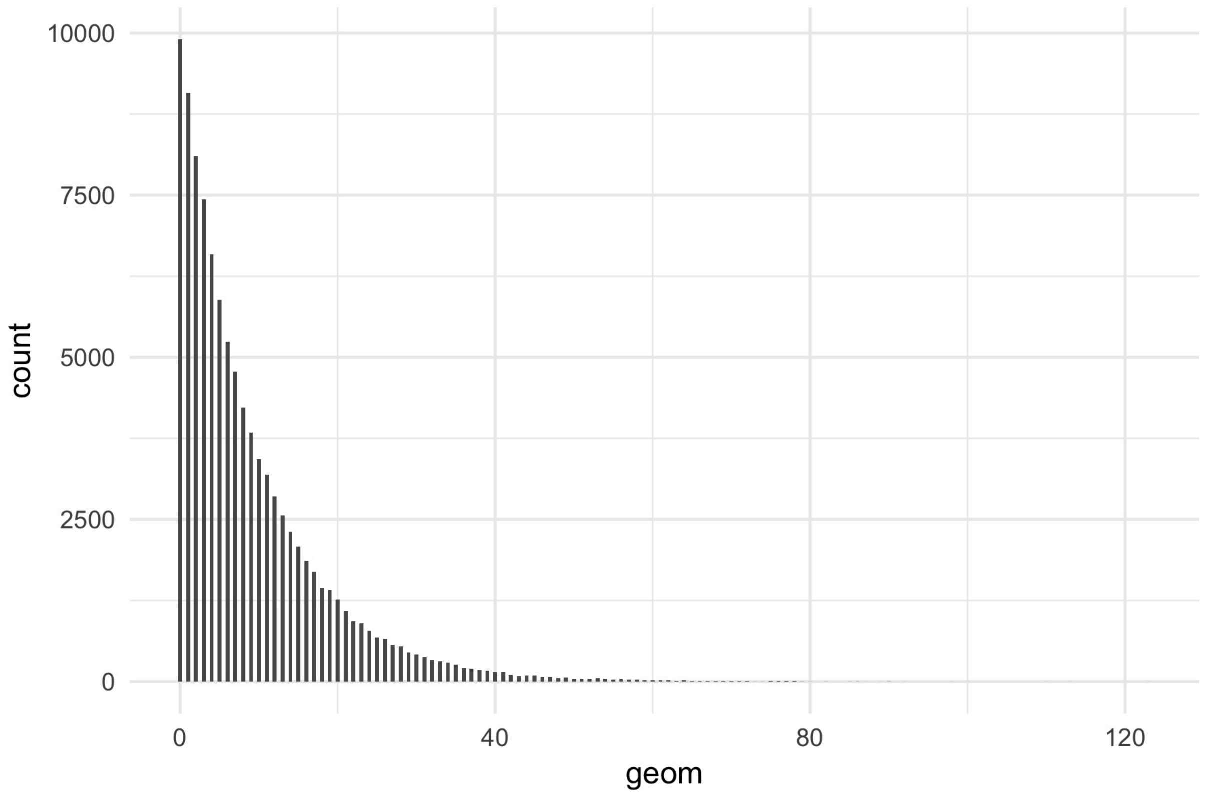
Simulating with rgeom

```
geom <- rgeom(100000, .1)
```

```
mean(geom)
# [1] 9.04376
```

$$X \sim \mathrm{Geom}(p)$$

$$E[X] = \frac{1}{p} - 1$$



FOUNDATIONS OF PROBABILITY IN R

Let's practice!