FOUNDATIONS OF PROBABILITY IN R

Probability of event A and event B

David Robinson
Data Scientist, Stack Overflow

Event A: "Coin is heads"

A = 1

A = 0

Events A and B: Two Different Coins

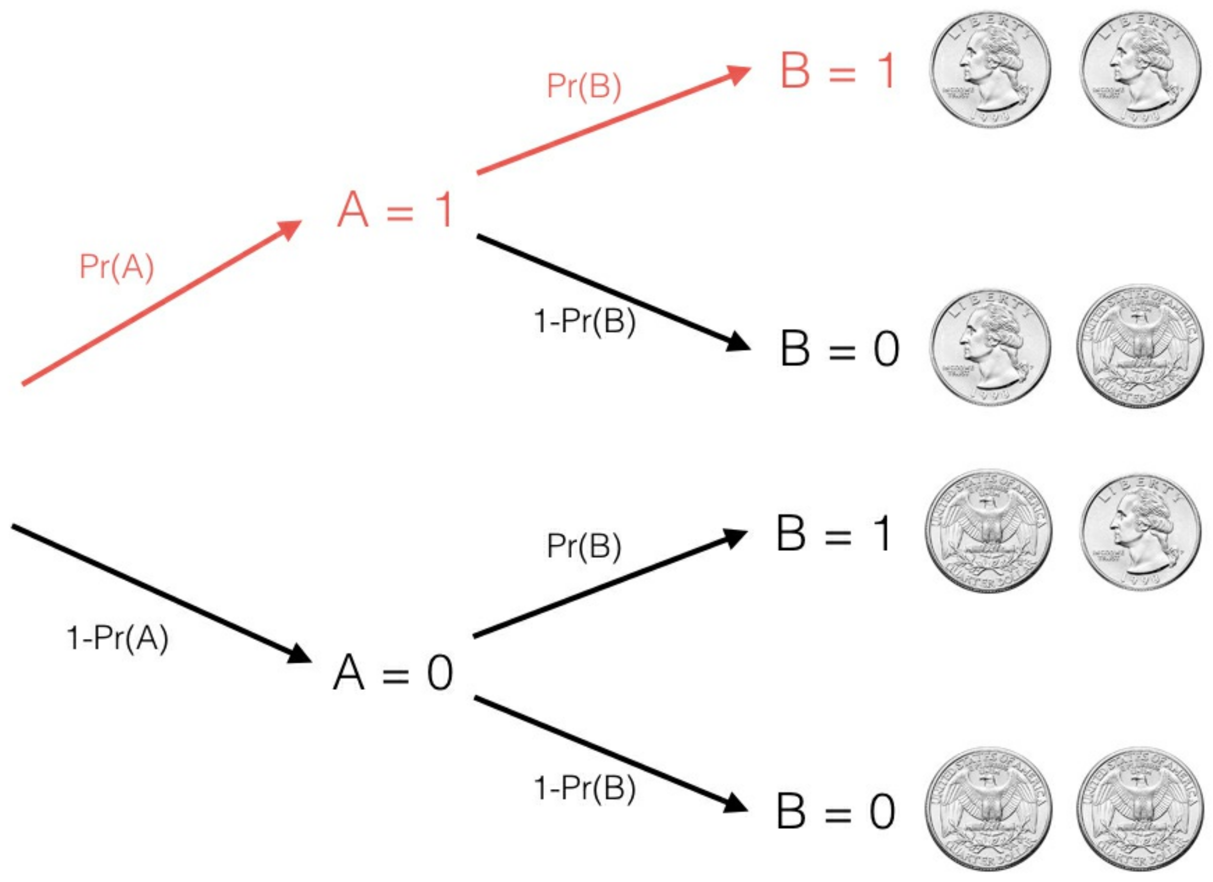
A = 1

A = 0

B = 1

B = 0

Probability of A and B



$$\Pr(A \text{ and } B) = \Pr(A) \cdot \Pr(B)$$

Simulating two coins

```r
A <- rbinom(100000, 1, .5)
```

```r
B <- rbinom(100000, 1, .5)
```

```r
A & B
# [1]  FALSE TRUE FALSE  FALSE...
```

```r
mean(A & B)
[1] 0.24959
```

$$\Pr(A \text{ and } B) = \Pr(A) \cdot \Pr(B)$$

$$\Pr(A \text{ and } B) = .5 \cdot .5 = .25$$

```r
A <- rbinom(100000, 1, .1)
```

```r
B <- rbinom(100000, 1, .7)
```

```r
A & B
# [1]  FALSE FALSE FALSE  FALSE...
```

```r
mean(A & B)
[1] 0.07043
```

$$\Pr(A \text{ and } B) = \Pr(A) \cdot \Pr(B)$$

$$\Pr(A \text{ and } B) = .1 \cdot .7 = .07$$

FOUNDATIONS OF PROBABILITY IN R

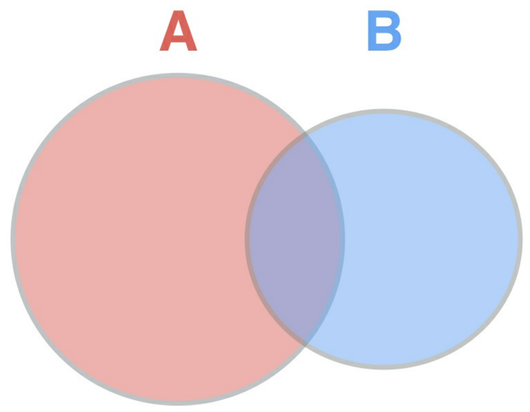
Let's practice!

FOUNDATIONS OF PROBABILITY IN R

Probability of A or B

David Robinson, PhD
Data Scientist, Stack Overflow

Probability of A or B



$$\Pr(A \text{ or } B) = \Pr(A) + \Pr(B) - \Pr(A \text{ and } B)$$

$$\Pr(A \text{ or } B) = \Pr(A) + \Pr(B) - \Pr(A) \cdot \Pr(B)$$

$$\Pr(A \text{ or } B) = .5 + .5 - .5 \cdot .5 = .75$$

Simulating two events

```r
A <- rbinom(100000, 1, .5)
```

```r
B <- rbinom(100000, 1, .5)
```

```r
mean(A | B)
[1] 0.75125
```

$$\Pr(A \text{ or } B) = \Pr(A) + \Pr(B) - \Pr(A \text{ and } B)$$

$$.75 = .5 + .5 - .5 \cdot .5$$

```r
A <- rbinom(100000, 1, .2)
```

```r
B <- rbinom(100000, 1, .6)
```

```r
mean(A | B)
[1] 0.6803
```

$$\Pr(A \text{ or } B) = \Pr(A) + \Pr(B) - \Pr(A \text{ and } B)$$

$$.68 = .2 + .6 - .2 \cdot .6$$

Three coins

$$\Pr(A \text{ or } B \text{ or } C)$$
$$= \Pr(A) + \Pr(B) + \Pr(C) -$$
$$\Pr(A \text{ and } B) - \Pr(A \text{ and } C) - \Pr(A \text{ and } B) +$$
$$\Pr(A \text{ and } B \text{ and } C)$$

```
mean(A | B | C)
```

FOUNDATIONS OF PROBABILITY IN R

Let's practice!

FOUNDATIONS OF PROBABILITY IN R

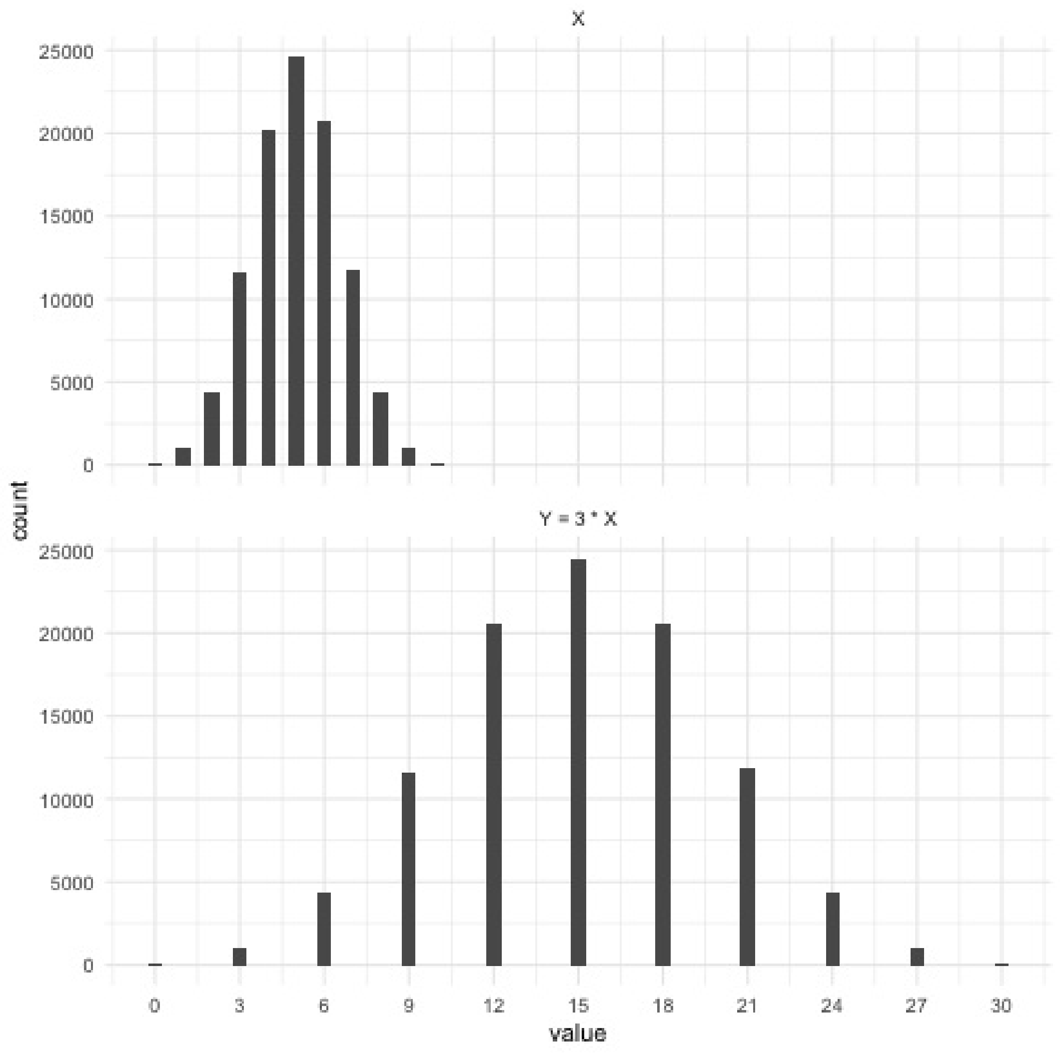
Multiplying random variables

David Robinson, PhD
Data Scientist, Stack Overflow

Multiplying a random variable

$$X \sim \text{Binomial}(10, .5)$$

$$Y \sim 3 \cdot X$$



Simulation: Effect of multiplying on expected value

$$X \sim \text{Binom}(10, .5)$$

$$Y = 3 \cdot X$$

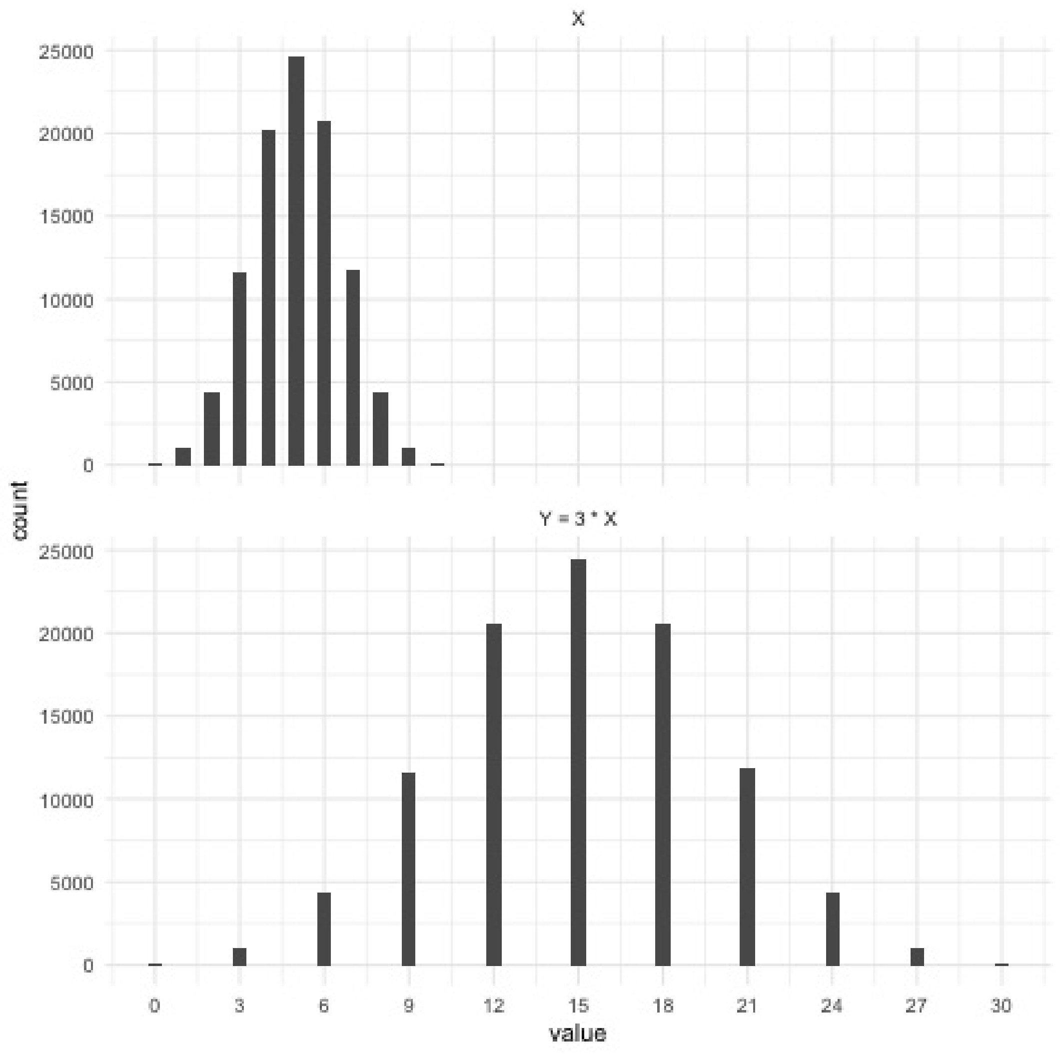
```
X <- rbinom(100000, 10, .5)
```

```
mean(X)
# [1] 5.006753
```

```
Y <- 3 * X
```

```
mean(Y)
# [1] 15.02026
```

$$E[k \cdot X] = k \cdot E[X]$$



Simulation: Effect of multiplying on variance

$$X \sim \text{Binom}(10, .5)$$

$$Y = 3 \cdot X$$

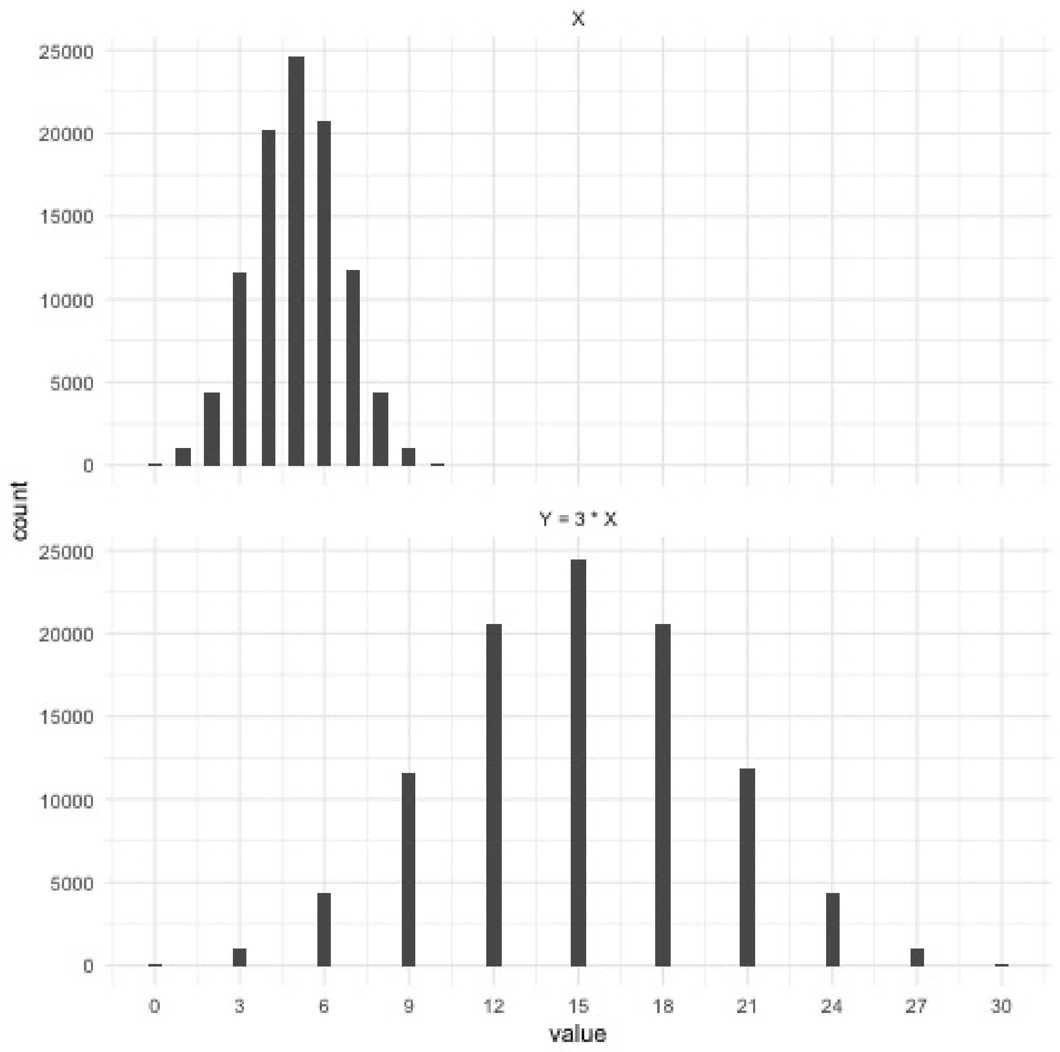
```
X <- rbinom(100000, 10, .5)
```

```
var(X)
# [1] 2.500388
```

```
Y <- 3 * X
```

```
var(Y)
# [1] 22.50349
```

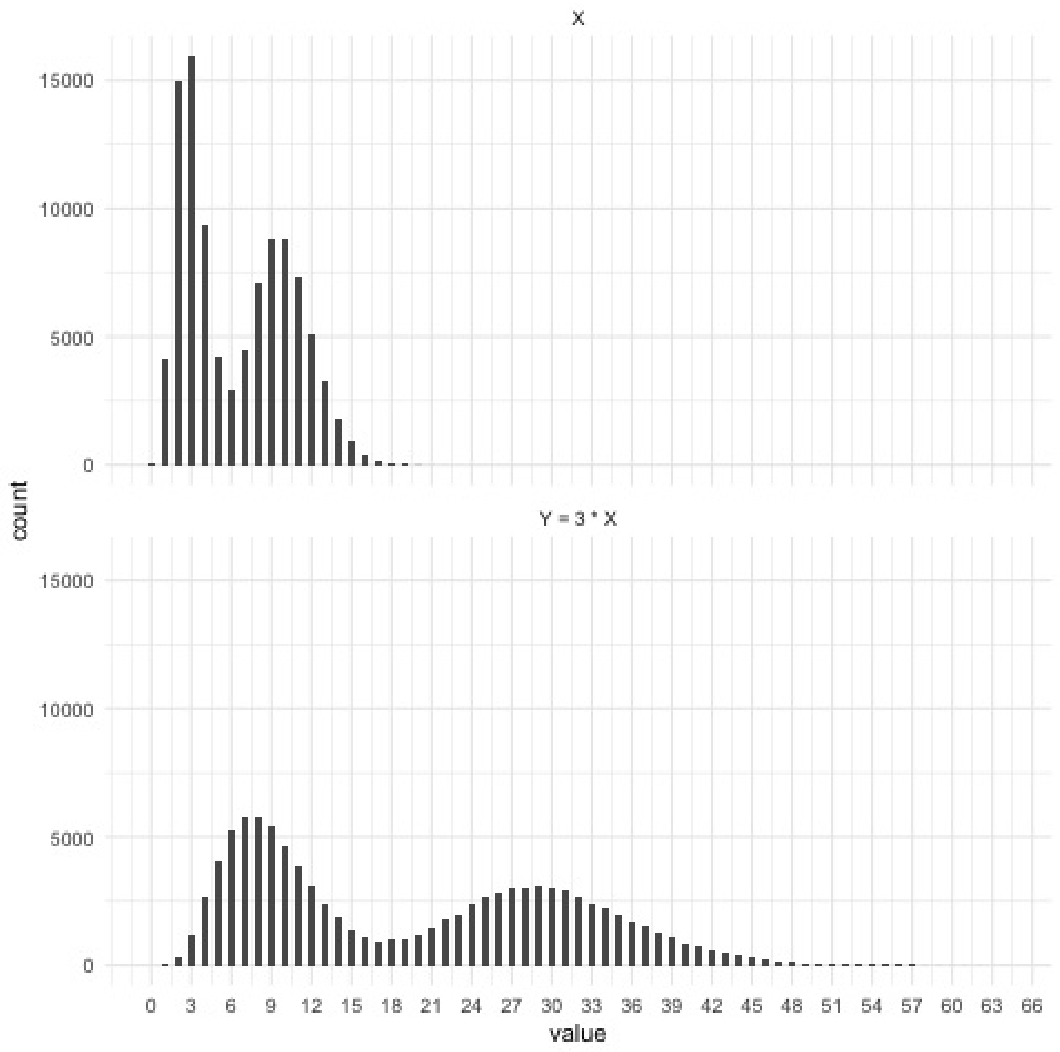
$$\text{Var}[k \cdot X] = k^2 \cdot \text{Var}[X]$$



Rules of manipulating random variables

$$E[k \cdot X] = k \cdot E[X]$$

$$\mathrm{Var}(k \cdot Y) = k^2 \cdot \mathrm{Var}(X)$$



FOUNDATIONS OF PROBABILITY IN R

Let's practice!

FOUNDATIONS OF PROBABILITY IN R

Adding two random variables together

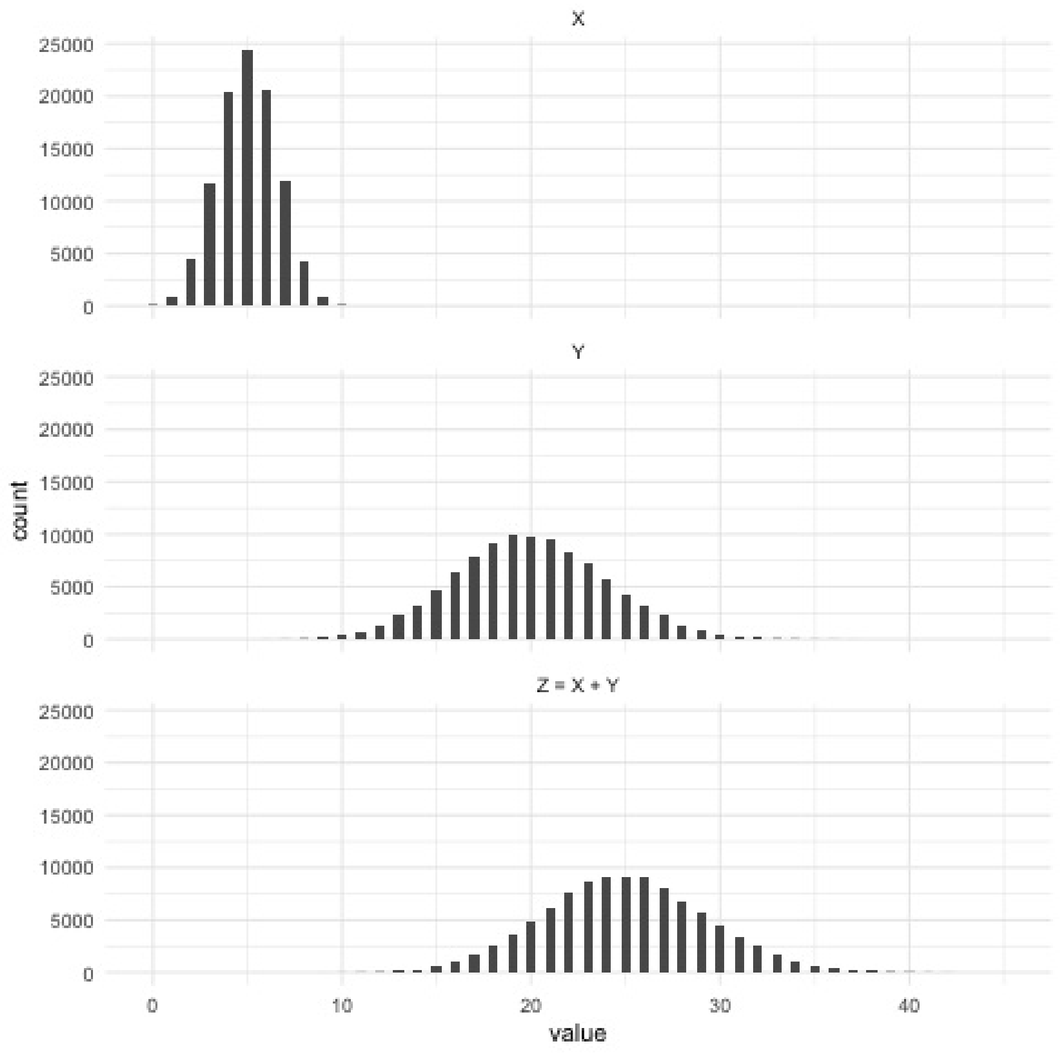
David Robinson
Data Scientist, Stack Overflow

Adding two random variables

$$X \sim \text{Binom}(10, .5)$$

$$Y \sim \text{Binom}(100, .2)$$

$$Z \sim X + Y$$



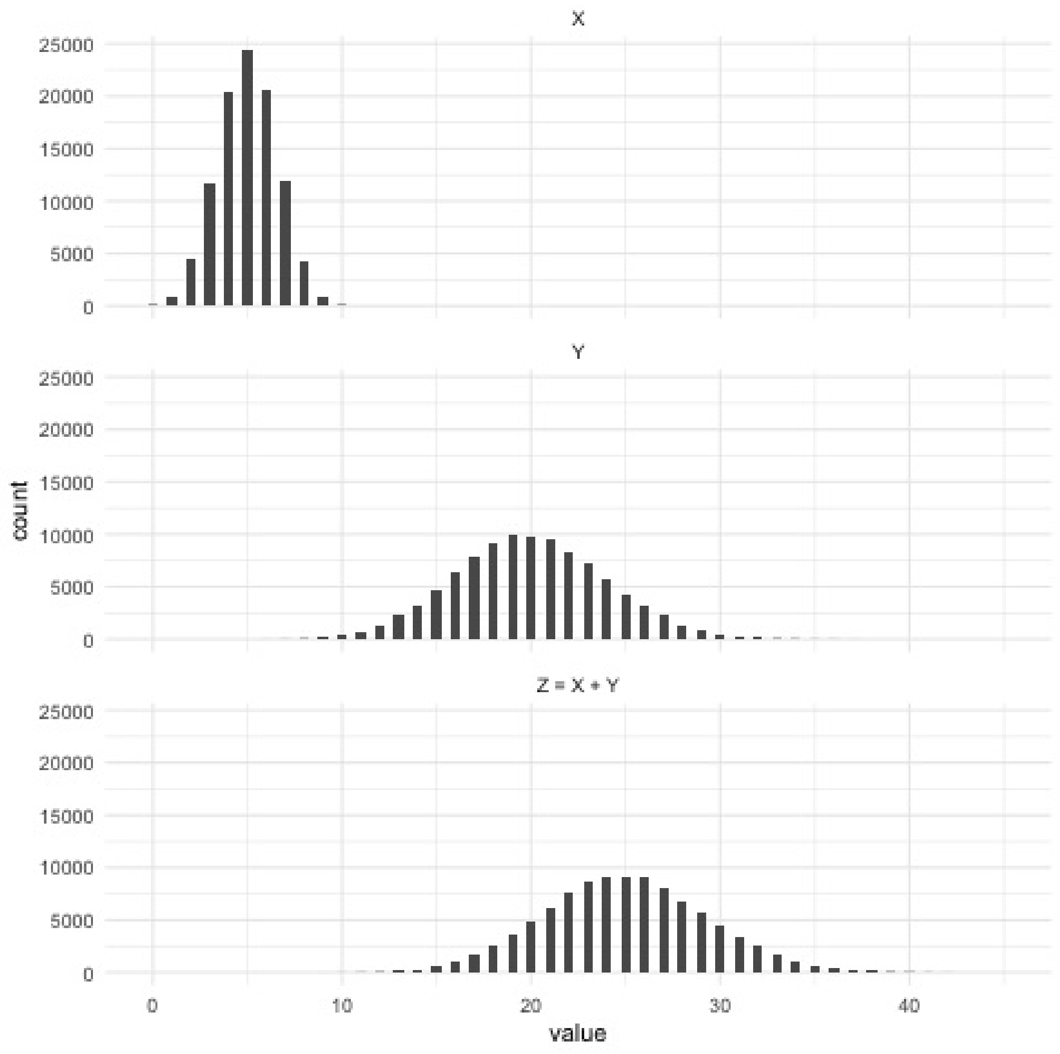
Simulation: expected value of X + Y

```
X <- rbinom(100000, 10, .5)
mean(X)
# [1] 5.00938
```

```
Y <- rbinom(100000, 100, .2)
mean(Y)
# [1] 19.99422
```

```
Z <- X + Y
mean(Z)
# [1] 25.0036
```

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



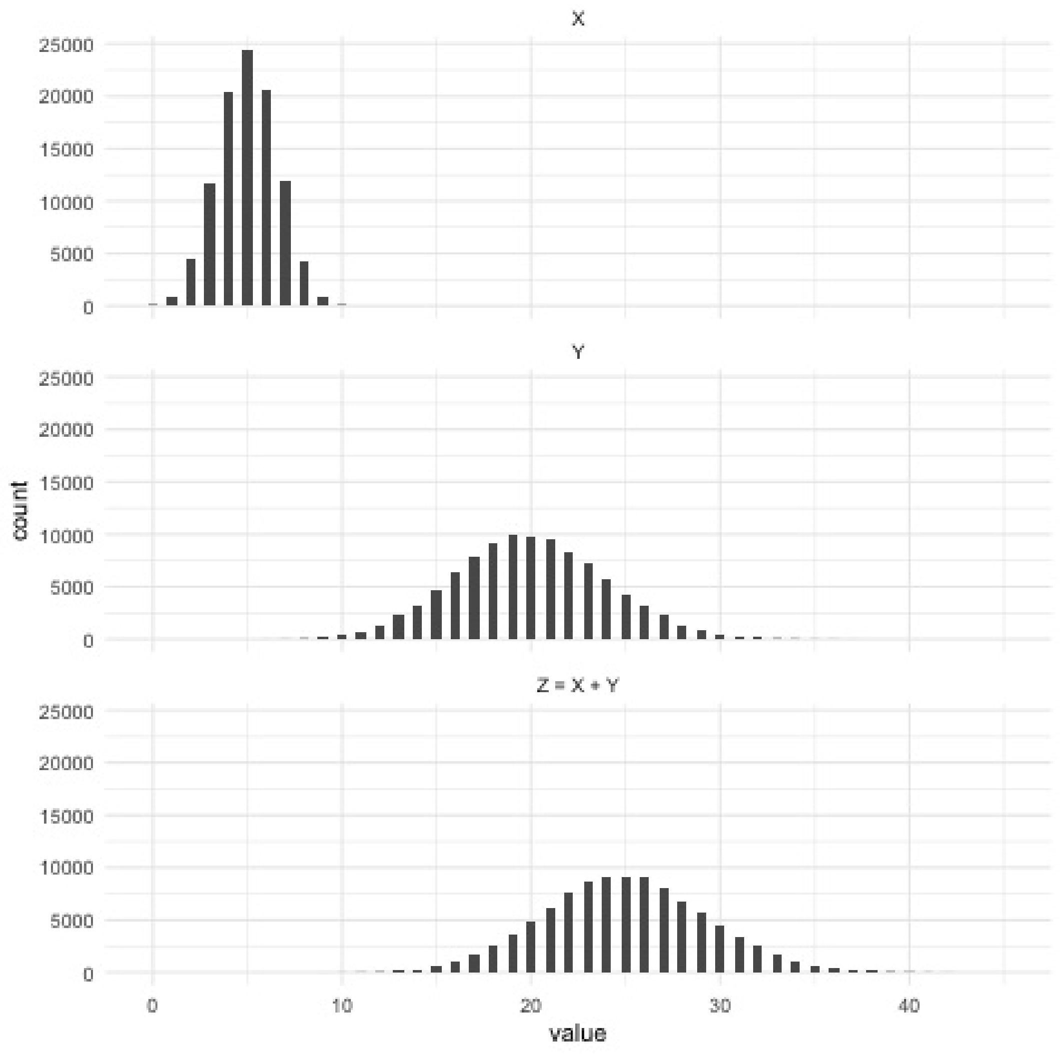
Simulation: variance of X + Y

```
X <- rbinom(100000, 10, .5)
var(X)
# [1] 2.500895
```

```
Y <- rbinom(100000, 100, .2)
var(Y)
# [1] 16.06289
```

```
Z <- X + Y
var(Z)
# [1] 18.58055
```

$$\text{Var}[X+Y] = \text{Var}[X] + \text{Var}[Y]$$



Rules for combining random variables

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

$$\text{(Even if X and Y aren't independent)}$$

$$\text{Var}[X+Y] = \text{Var}[X] + \text{Var}[Y]$$

$$\text{(Only if X and Y are independent)}$$

FOUNDATIONS OF PROBABILITY IN R

Let's practice!