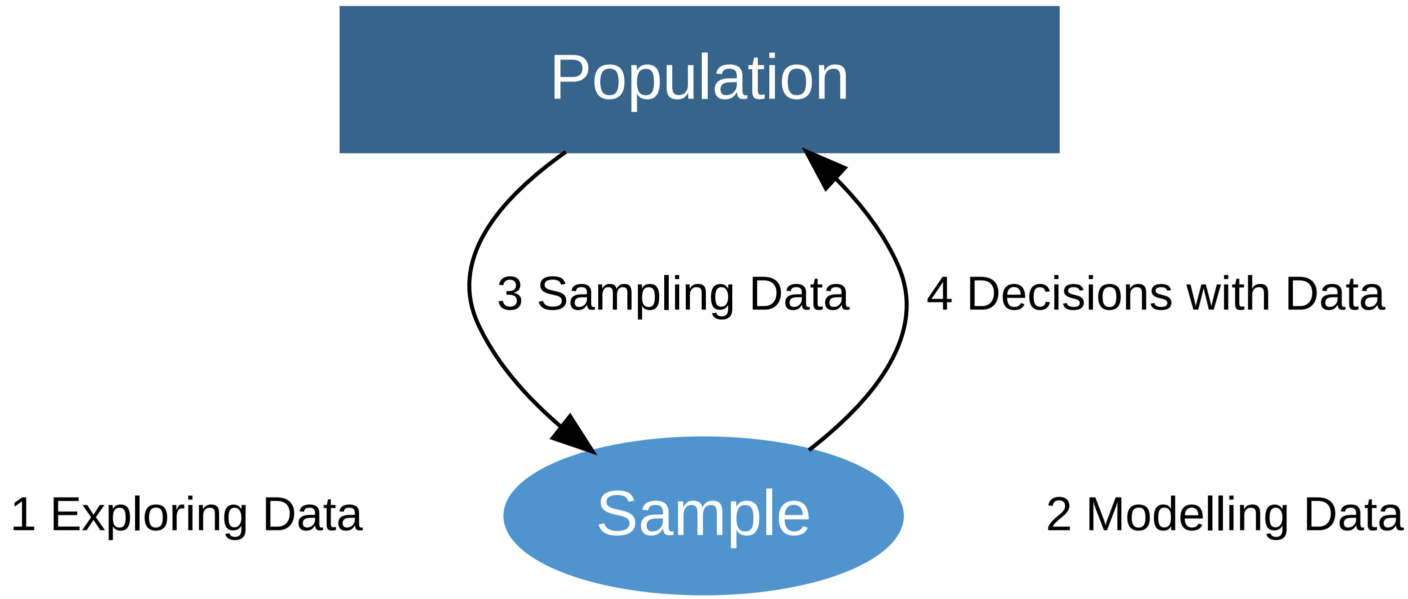
Binomial Formula

Sampling Data | Understanding Chance

Unit Overview



Module3 Sampling Data

Understanding Chance

What is chance?

Chance Variability

How can we model chance variability by a box model?

Sample Surveys

How can we model the chance variability in sample surveys?

Binomial Formula

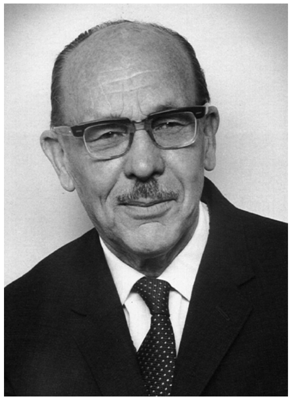
Data Story

Coin tossing in WWII

Coin tossing in WWII

- John Edmund Kerrick (1903–1985) was a mathematician noted for a series of experiments in probability which he conducted while interned in Nazi-occupied Denmark (Viborg, Midtjylland) in the 1940s.

- Kerrich had travelled from South Africa to visit his in-laws in Copenhagen, and arrived just 2 days after Denmark was invaded by Nazi Germany!



- Fortunately Kerrich was imprisoned in a camp in Jutland run by the Danish Government in a `truly admirable way'.

- With a fellow internee Eric Christensen, Kerrich set up a sequence of experiments demonstrating the empirical validity of a number of fundamental laws of probability.
    - They tossed a coin 10,000 times and counted the number of heads.
    - They made 5000 draws from a container with 4 ping pong balls (2x2 different brands), ‘at the rate of 400 an hour, with - need it be stated - periods of rest between successive hours.’
    - They investigated tosses of a “biased coin”, made from a wooden disk partly coated in lead.
- In 1946 Kerrich published his finding in a monograph An Experimental Introduction to the Theory of Probability.

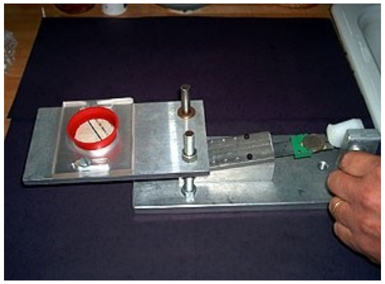
Statistical Thinking

Kerrich and Christensen tossed a coin 10,000 times. How many heads do you think they counted?

Famous Coin tossers

Famous Coin tossers

Person	Number of Tosses	Number of Heads	P(Head)
Count Buffon (1707-1788)	4040	2048	0.507
Karl Pearson (1857-1936)	24000	12012	0.5005
John Kerrick (1903-1985 war camp)	10000	5067	0.5067
Perci Diaconis (1945-present)	machine		0.51 (if started head)



Simulating their tosses

```
set.seed(5)
table(sample(c("H","T"),4040,replace=T))/4040
```

```
##
##         H         T
## 0.5111386 0.4888614
```

```
table(sample(c("H","T"),24000,T))/24000
```

```
##
##         H         T
## 0.5024583 0.4975417
```

```r
table(sample(c("H","T"),10000,T))/10000
```

```
##
##      H      T
## 0.4987 0.5013
```

Binomial coefficients

Binomial coefficients

We can work out the exact chance of P(Head) by using the Binomial model.

Factorial

The number of ways of rearranging n distinct objects in a row is **n factorial**:

- n! = nx(n-1)x(n-2) … 3x2x2
- 0! = 1

Example

- Show 6! = 720

```r
factorial(6)
```

```
## [1] 720
```

```r
x=c(1:6)
prod(x)
```

```
## [1] 720
```

Binomial coefficient

- Suppose we have n objects in a row, made up of 2 types: x (type1) and $n - x$ (type2).

- The number of ways of rearranging the n objects is given by the **binomial coefficient**:

$$\binom{n}{x} = \frac{n!}{x!(n-x)!}$$

Special case: $\binom{n}{0} = 1$

Example

- Show $\binom{6}{2} = 15$
- Give an example of what it could represent?

```r
choose(6,2)
```

```
## [1] 15
```

Binomial Model

Binary trials

Binary trials

A **binary trial** is where only 2 things can occur, so P(event) = p and P(not event) = 1-p.

What are some examples?

Binomial Theorem

Binomial theorem

- Suppose we have n independent, binary trials, with P(event)=p at every trial, and n is fixed.
- The chance that exactly x events occur is

$$\binom{n}{x} p^x (1-p)^{n-x}$$

Example

- A fair coin is tossed 5 times. What is the probability of getting 3 heads?
- A fair coin is tossed 500 times. What is the probability of getting 300 heads?

```
dbinom(3,5,0.5)    #dbinom(k,n,p) = P(exactly k events)
```

```
## [1] 0.3125
```

```
dbinom(300,500,0.5)
```

```
## [1] 1.544255e-06
```

Example

Suppose 100 babies are born at RPA hospital today with $P(boy) = 0.51$.

- What is the probability that 55 boys were born?
- What are your assumptions?

```r
dbinom(55,100,0.51)
```

```
## [1] 0.05803559
```

Justifying the Binomial Theorem

- Suppose we have 4 independent, binary trials, with P(event)=p at every trial.
- To get exactly 3 events occuring we can have the following sequences:

(E E E NE), (E E NE E), (E NE E E), (NE E E E)

- There are $\binom{4}{3}$ possible sequences.
- Each sequence has a chance $p^3(1-p)$ of occuring.
- So the final answer is $\binom{4}{3}p^3(1-p)$.

Summary

binomial coefficient, factorial, binomial formula

Key Words

Further Thinking