Lecture 3: Binomial Distribution

Sta 111

Colin Rundel

May 16, 2014

Ordering a deck of cards ...

If you have ever shuffled a deck of cards you have done something no one else has ever done before or will ever do again ...

There are approximately 8×10^{67} (52!) possible configurations of a deck of 52 to cards

To put that in context:

- Cells in the human body (10^{14})
- Seconds since the big bag (10^{18})
- Grains of sand on all beaches on earch (7.5×10^{18})
- Stars in the universe (10^{23})
- Atoms in the observable universe (10^{80})
- A Googol (10^{100})

Example

Imagine you have a bag with 6 slips of paper numbered 1 to 6. How many different pairs can you draw if you sample without replacement?

Permutations & Combinations

If we have n items and want to select k of them without replacement, then there are j possible outcomes.

Permutations - when we care about the order in which pull out the items:

$$j = \frac{n!}{(n-k)!}$$

Combination - when we *do not* care about the order in which pull out the items:

$$j = \binom{n}{k} = \frac{n!}{(n-k)!\;k!}$$

Permutations & Combinations - Derivation

Some properties of the Binomial coefficient

$$\binom{n}{k} = \binom{n}{n-k}$$

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n$$

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}, \text{ for } 0 < k < n$$

Pascal's Triangle and the Binomial coefficient

$$\binom{0}{0}$$

$$\binom{1}{0} \quad \binom{1}{1}$$

$$\binom{2}{0} \quad \binom{2}{1} \quad \binom{2}{2}$$

$$\binom{3}{0} \quad \binom{3}{1} \quad \binom{3}{2} \quad \binom{3}{3}$$

$$\binom{4}{0} \quad \binom{4}{1} \quad \binom{4}{2} \quad \binom{4}{3} \quad \binom{4}{4}$$

$$\vdots$$

Probability Distributions

Description of the probability for all values of a random variable (a numeric value that depends in some way on chance).

We have to distinguish between the discrete and continuous case:

- Discrete (integer valued) - easy to assign probability to each event (even if there are infinitely many)
  - Value of the role of a six sided die
  - Number of coin flips until the first head
- Continuous (real valued) - probability defined based on an interval
  - Probability a student's height is *exactly* 5'9"
  - Probability a students height is between 5'9" and 5'10"

Distributions Functions

If X takes countably-many values, then we can list them and just report

$$f(x) = P(X = x)$$

- When X is discrete - Probability *mass* function

$$P(X \in A) = \sum_{x \in A} f(x)$$

- When X is continuous - Probability *density* function

$$P(X \in A) = \int_A f(x)dx$$

Probability mass function

Let X be the number of aces in two draws *without replacement* from a 52-card deck.

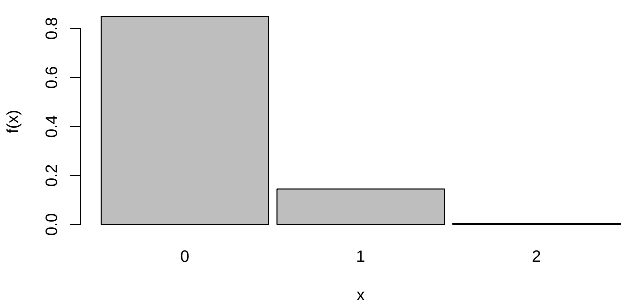
The PMF can be presented in a variety of ways:

	x	f(x)
No aces	0	$(48/52)(47/51) = 0.8507$
One ace	1	$2(4/52)(48/51) = 0.1448$
Two aces	2	$(4/52)(3/51) = 0.0045$

Probability mass function

Let X be the number of aces in two draws *without replacement* from a 52-card deck.

The PMF can be presented in a variety of ways:



Probability mass function

Let X be the number of aces in two draws *without replacement* from a 52-card deck.

The PMF can be presented in a variety of ways:

For $x \in \{0, 1, 2\}$,

$$P(X = x) = \frac{\binom{4}{x}\binom{48}{2-x}}{\binom{52}{2}} = \frac{4!}{x!(4-x)!}\,\frac{48!}{(2-x)!(46+x)!}\,\frac{2!\;50!}{52!}$$

$$\approx \frac{2.25 \times 10^{59}}{x!(4-x)!(2-x)!(46+x)!}$$

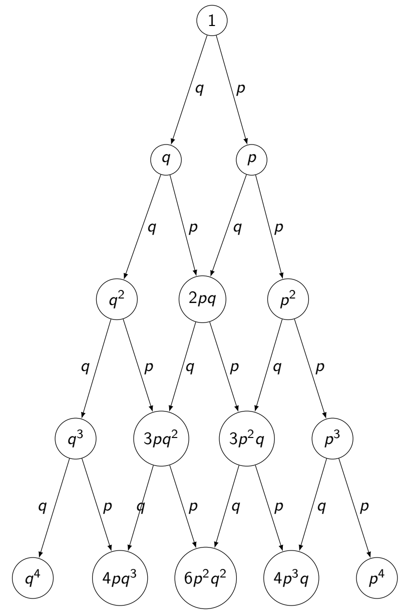
Example

A company is testing a new manufacturing process for aluminum cases, if 20% of the cases do not meet their specifications what is the probability that if the company checks the next four cases that only one of them will not meet specification?

Let the probability a test succeeds be $p = 0.8$ and the probability a test fails be $q = 0.2$ then

$$\begin{aligned} P(\text{1 failure in 4 tests}) &= pppq + ppqp + pqpp + qppp \\ &= 4p^3q \\ &= \binom{4}{1} p^3 q \end{aligned}$$

Binomial Distribution



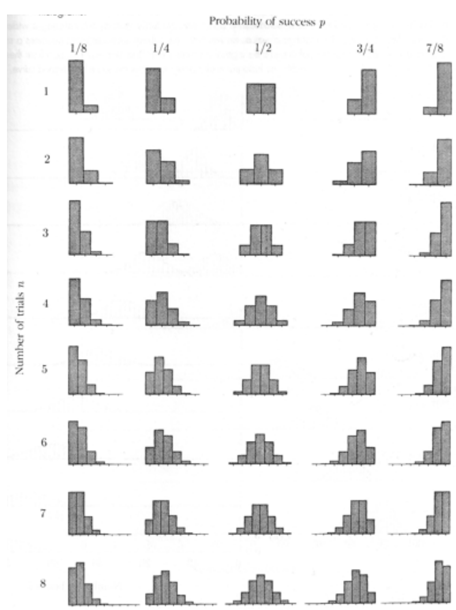
Binomial Distribution

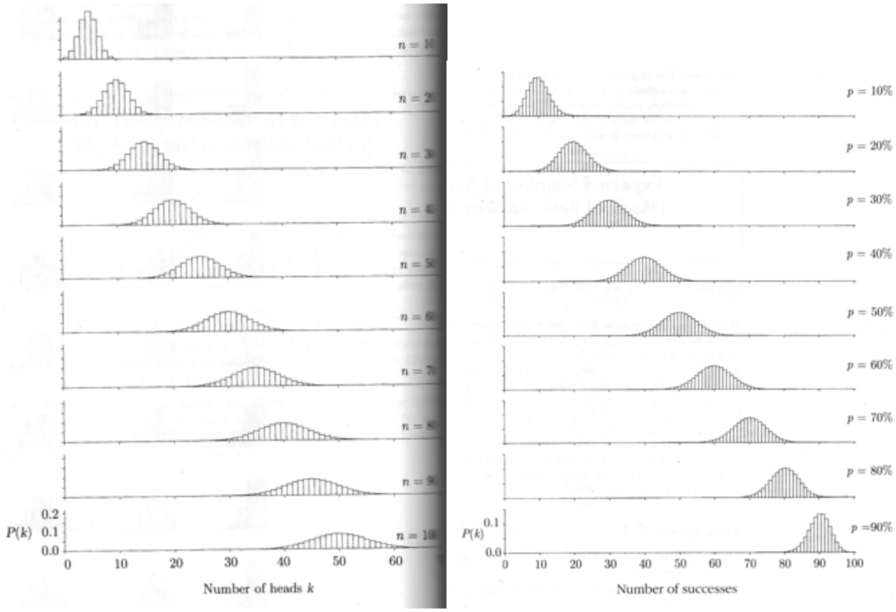
We define a random variable X that reflects the *number of successes* in a *fixed number* of *independent trials* with the *same probability of success* as having a binomial distribution.

If there are n trials then

$$X \sim \text{Binom}(n, p)$$

$$f(k|n,p) = P(X = k|n,p) = \binom{n}{k} p^k (1-p)^{n-k}$$





What is the most probable outcome?

The binomial distribution is unimodal which makes our life easier...

We can look at the ratio of successive outcomes,

$$r = \frac{P(X = k+1)}{P(X = k)}, \text{ for } 0 \le k \le n-1$$

r is largest when $k = 0$ and gets progressively smaller.

When $r > 1$ then $P(X = k+1) > P(X = k)$
When $r < 1$ then $P(X = k+1) < P(X = k)$

Maximum (mode) of the distribution occurs when r switches from being greater than 1 to less than 1.

What is the most probable outcome? cont.

What value of k results in $r \le 1$?

What is the most probable outcome? cont.

Max probability is therefore the smallest integer value of $k \ge np - q$.

We can narrow that relationship down somewhat since there must be an integer value of k between

$$\begin{aligned} np - q \le k &\le np - q + 1 \\ np - q \le k &\le np - q + (p+q) \\ np - q \le k &\le np + p \end{aligned}$$

Special case when $r = 1$ as it implies that $P(X = k) = P(X = k+1)$ in which case both values are equally probable.

What is the scale of this maximum probability?

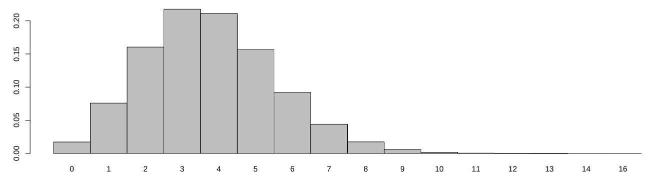
Not very large...

$P(X = k)$ maxes out at a little bit less than $1/\sqrt{n}$, therefore $P(X = k_{\text{mode}}) \to 0$ as $n \to \infty$.

Conceptually, as the number of bins increases the mass in each bin must necessarily get smaller, we are in essense moving from discrete to continuous distribution.

Some examples...

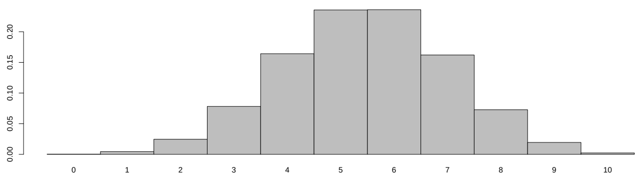
Let $X \sim \text{Binom}(25, 0.15)$ then the distribution of X looks like



$$k_{\text{mode}} = \lfloor (np - q, np + p) \rceil = \lfloor (2.9, 3.9) \rceil = 3$$

Some examples...

Let $X \sim \text{Binom}(10, 6/11)$ then the distribution of X looks like



$$k_{\text{mode}} = \lfloor (np - q, np + p) \rceil = \lfloor (5, 6) \rceil = 5, 6$$

Outcome Ranges

Often it is more interesting to talk about the probability of a range of outcomes.

For example, going back to the manufacturing example (where p=0.8, n=4) what is the probability that there are 1 or fewer defective cases?

$$\begin{aligned}
P(\text{1 or fewer defective}) &= P(\text{3 or more successes}) \\
&= P(X = 3 \text{ or } 4) \\
&= P(X = 3) + P(X = 4) \\
&= \binom{4}{3}(0.8)^3(0.2)^1 + \binom{4}{4}(0.8)^4(0.2)^0 \\
&= 4(0.1024) + 1(0.4096) \\
&= 0.8192
\end{aligned}$$

What happens for large n?

Let $x = z + np$ and $c = P(np)$ then,

$$
\begin{aligned}
\log P(np+z) &\approx \log P(np) - \frac{z^2}{2npq} \\
\log P(x) &\approx \log P(np) - \frac{(x-np)^2}{2npq} \\
P(x) &\approx \exp\left(\log P(np) - \frac{1}{2}\frac{(x-np)^2}{npq}\right) \\
P(x) &\approx c\ e^{-\frac{1}{2}\frac{(x-np)^2}{npq}}
\end{aligned}
$$

de Moivre-Laplace Limit Theorem

When n is large enough the Binomial distribution will always have this bell-curve shape.

Shape of the curve given by $c\ e^{-b(x-a)^2}$

de Moivre and Laplace where the first to identify this pattern and characterize the shape of the curve (by finding a, b, c).

This is a special case of a more general result known as the Central Limit Theorem. (More on this later)