Scientific Measurement

PHY-103

Dr. Eram Rizvi & Dr. Jeanne Wilson

Lecture 5 - The Binomial Probability Distribution

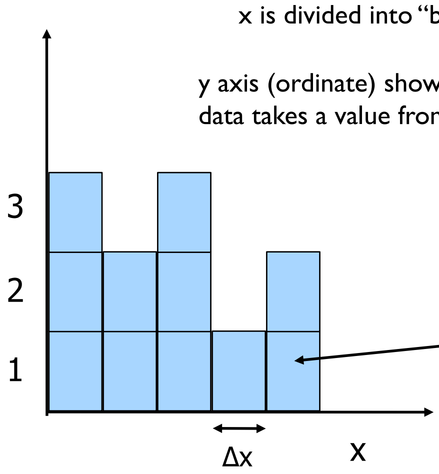
Queen Mary
University of London

Binomial Distribution

- We often want to know likelihood of an event occurring
- Quantify this with dimensionless parameter: Probability (0→1)
- Important concept in e.g. quantum physics
- Often use histograms to visualise this

x is divided into "bins" of width Δx

y axis (ordinate) shows number of times your data takes a value from x → x+Δx



Each x bin i contains a number N_i

In fact it is N_i divided by bin width Δx

$N_i/\Delta x$

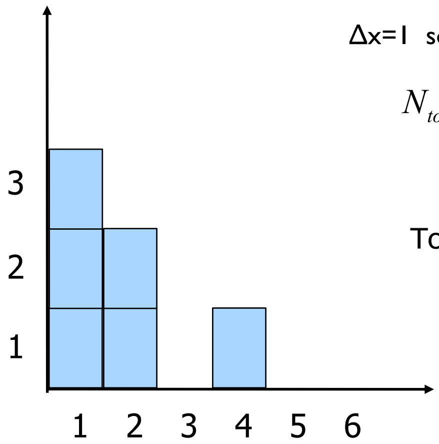
Dividing by N_{tot} each bin contains $\dfrac{N_i/N_{tot}}{\Delta x}$

i.e. the fraction of data in region x → x+Δx

Thus histogram becomes a probability density since $\sum_i \dfrac{N_i/N_{tot}}{\Delta x} \times \Delta x = \sum_i N_i/N_{tot} = 1$

Thus a normalised histogram is a probability density!

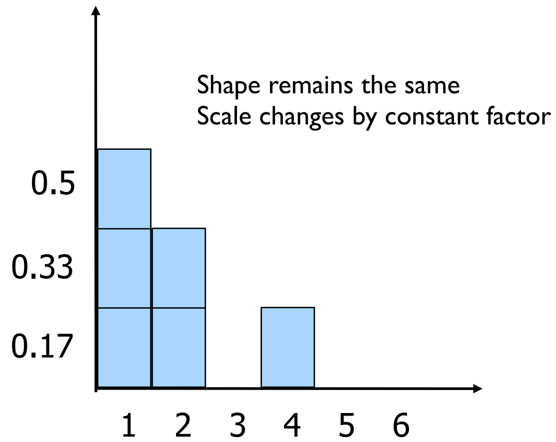
Consider die throwing experiment. Throw die 6 times and histogram the results



$\Delta x = 1$ so $N_i/\Delta x = N_i$

$$N_{tot} = \sum_i N_i = 6$$

$$\text{Total Probability} = \sum_i \frac{N_i / N_{tot}}{\Delta x} \times \Delta x = \sum_i N_i/6 = 1$$



Now let $\Delta x \rightarrow 0$

Histogram → continuous distribution (for non-discrete observables)

compare dice throwing result to prob of bus arriving in time Δt

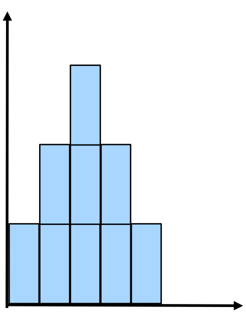
Let's halve the bin width:

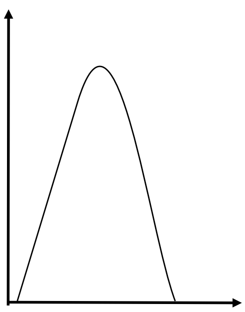
$$\text{Bin content} \approx \frac{\frac{1}{2} N_i / N_{tot}}{\frac{1}{2}\Delta x} \quad \text{i.e. it is unchanged!}$$

Given enough data:

$\Delta x \rightarrow dx$

$N_i/\Delta x \rightarrow f(x)$





Normalisation condition is then $\int_{-\infty}^{+\infty} f(x)\,dx = 1$

Prob of bus arriving in time dt = 1ns is tiny, but summed over 10^{11} 'bins' (~1 min) prob becomes sensible value

- Lets look at the discrete case of a dice throwing experiment
- Calculate probability of different outcomes
- Result is a discrete variable
  - a die can only result in 1,2,3...6 i.e. not 2.45
  - a coin toss only results in heads or tails (never h and t!)
- Consider 6 tosses of a coin (equivalent to single toss of 6 coins)
- How often does this result in 6 heads (6h)?
- P(6h) = $(½)^6$ = 1/64 = 1.6%
- Repeat this experiment 100 times:
  we expect 6h to happen 1 or 2 times (maybe 0 or 3 times)

- What is prob of 1 head & 5 tails i.e. P(1h) ?
  $(½)^1$ x $(½)^5$? NO! This is P(htttttt) what about P(thtttt) etc?
  P(1h) = 6$(½)^1$ x $(½)^5$ allows h on any of the 6 tosses
  P(1h) = 6/64 = 9.4% Quite likely to occur in 100 identical experiments

- What about P(4h) ?
  P(4h) = $(½)^4$ x $(½)^2$ x N
  N = Number of combinations of getting 4 out of 6

No. of h	N
0	1
1	6
2	15
3	20
4	15
5	6
6	1

h h h h t t
h h h t h t
h h t h h t
...

h h h t t t
h h t h t t
h t h h t t
...

4 h is like 2 h etc...

Easier way to determine N: ${}_nC_k$ = number of combinations of k from n

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

Number of ways of picking 4 heads from 6 tosses is:

$$ {}_{6}C_{4} = \frac{6!}{(6-4)!4!} = \frac{6\times 5}{2} = 15 $$

$$ \mathrm{P(4h)} = {}_{6}C_{4}\left(\frac{1}{2}\right)^{4}\left(\frac{1}{2}\right)^{2} = \frac{15}{64} = 0.234 \qquad (=23.4\%) $$

We took P(h) = P(t) = ½

i.e. P(success) = P(fail)

What if P(success) ≠ P(fail) ?

Let p = P(success)

Binomial Distribution

$$ \boxed{f(k;n,p) = \frac{n!}{(n-k)!k!} p^{k}(1-p)^{n-k}} $$

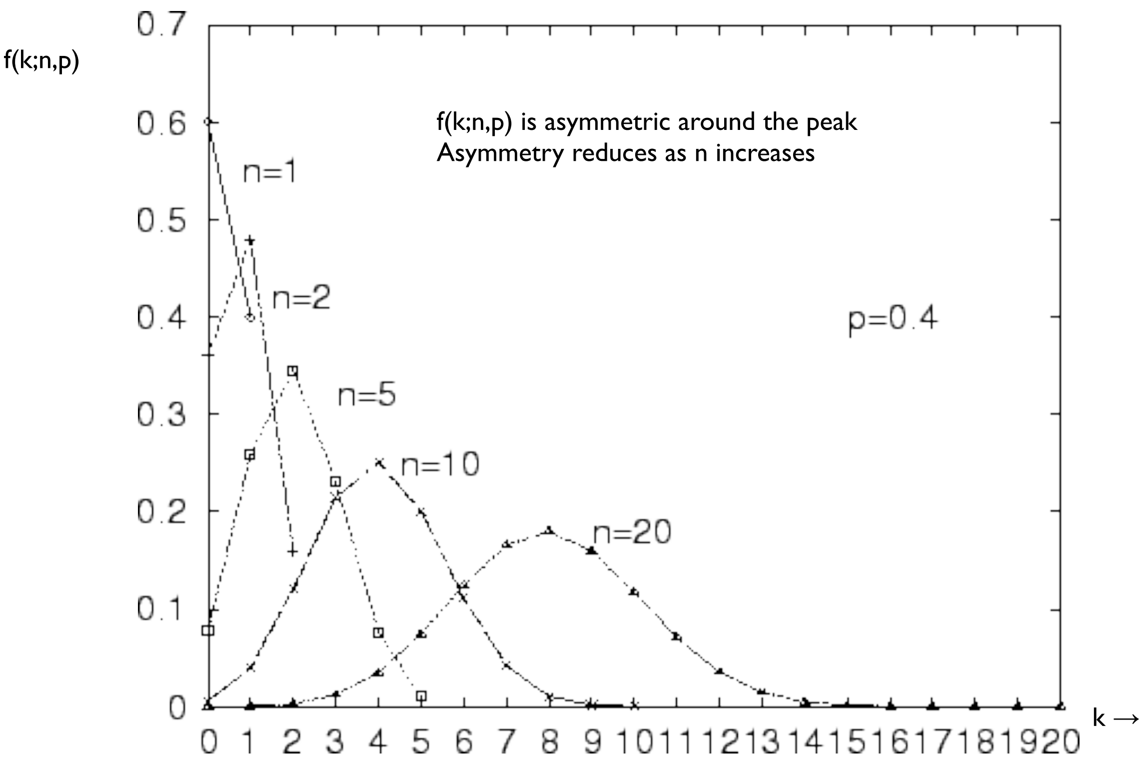
$$ = {}_{n}C_{k}\, p^{k}(1-p)^{n-k} $$

$$ = {}_{n}C_{k}\, p^{k} q^{n-k} \qquad \text{where } q = 1-p \qquad \text{q=P(fail)} $$

Prob. of k successes given n trials,
and probability of success for single trial = p

Binomial distribution for various n values with p=0.4

Note the distribution is <u>discrete</u> - lines between points are to guide the eye



- Binomial distribution describes statistics of situations with a true/false outcome only
- Distributions has some nice properties:

$$\bar{k} = \text{mean number of successes} = \sum_{k=0}^{n} k \times f(k) = np$$

i.e. mean number of heads from 6 tosses = 3

standard deviation: $$\sigma_k = \sqrt{np(1-p)}$$
$$= \sqrt{npq}$$

- In general p ≠ 0.5 e.g. Probability(England winning next match)
- Binomial statistics used for true/false type experiments
- Example: opinion polls

Ask a sample of people a question. Allow only yes/no answers

"Would you prefer to see the re-introduction of the death penalty to the UK?"

Poll 1000 randomly selected people

n = 1000

Yes = 500 No (or "Don't know") = 500

Thus p = 500/1000 = 50% (!)

What is the uncertainty?

$$\sigma = \sqrt{np(1-p)}$$
$$= \sqrt{1000 \times 0.5 \times 0.5}$$
$$= \sqrt{250} = 16$$

$$\text{Fractional error} = \frac{\sigma}{np} = \frac{16}{500} \approx 3\%$$

- To reduce this error by factor of three to 1% need to work very hard

$$\frac{\sqrt{n \times 0.50 \times 0.50}}{n \times 0.50} = 1\% = \frac{1}{\sqrt{n}}$$

$$\sqrt{n} = \frac{1}{0.01}$$

$$n = 10000$$

Need to increase n by factor of 10 for a factor $\sqrt{10}$ reduction in error!