Part IA Engineering Mathematics: Lent Term

Convolution

Fourier Series

Probability

Joan Lasenby

Lent 2014

Section 9

Probability

In this section we summarise the key issues in pages 1–13 of the basic probability teach-yourself document and provide a single simple example of each concept.

This presentation is intended to be reinforced by the many examples in the teach-yourself document and the first 12 questions of examples paper 10.

Probability

$$\text{Probability of } A \; = \; \frac{\text{Number of outcomes for which } A \text{ happens}}{\text{Total number of outcomes (sample space)}}$$

What is the probability of drawing an ace from a shuffled pack of cards? There are 4 aces. There are 52 cards in total. Therefore the probability is

$$\mathbf{P}(\text{ ace }) = \frac{4}{52} = \frac{1}{13}$$

Adding Probabilities

$$\mathbf{P}(A \text{ or } B) = \mathbf{P}(A) + \mathbf{P}(B)$$

provided A and B cannot happen together, i.e. A and B must be mutually *exclusive* outcomes.

What is the probability of drawing an ace or a king from a shuffled pack of cards?

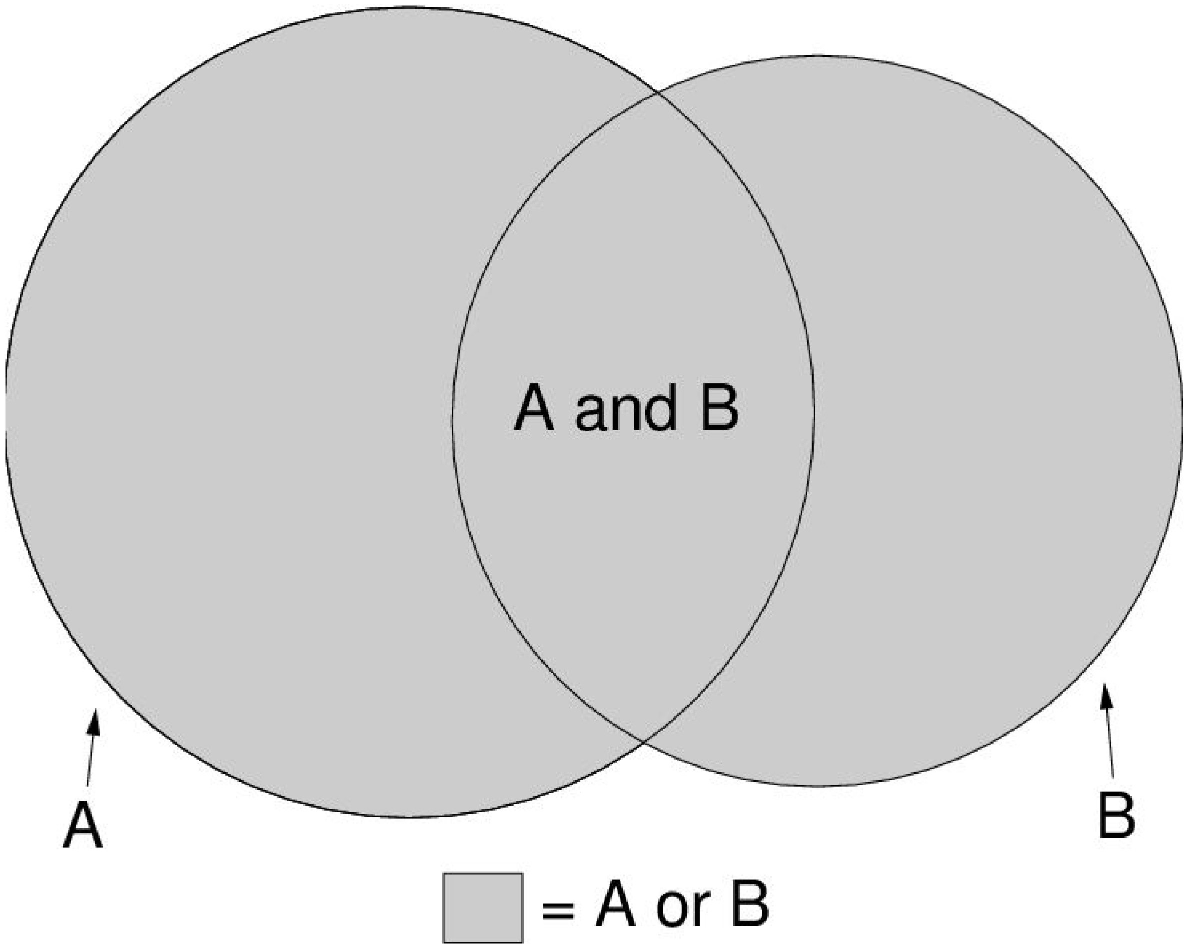
$$\mathbf{P}(\text{ ace }) = \frac{1}{13}$$

$$\mathbf{P}(\text{ king }) = \frac{1}{13}$$

$$\Rightarrow \mathbf{P}(\text{ ace or king }) = \frac{1}{13} + \frac{1}{13} = \frac{2}{13}$$

When not to add Probabilities

When the events are not mutually exclusive.



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

Non-Exclusive Events

What is the probability of drawing an ace or a spade from a shuffled pack of cards?

$$\mathbf{P}(\text{ ace }) = \frac{1}{13} \qquad \mathbf{P}(\text{ spade }) = \frac{13}{52} = \frac{1}{4}$$

but $\mathbf{P}(\text{ ace or spade })$ is not the sum of these values because the outcomes “ace” and “spade” are not exclusive; it is possible to have them both together by drawing the ace of spades.

To calculate $\mathbf{P}(\text{ ace or spade })$

either use the formula from the previous slide:

$$\mathbf{P}(\text{ ace }) + \mathbf{P}(\text{ spade }) - \mathbf{P}(\text{ ace of spades })$$

$$\frac{1}{13} + \frac{1}{4} - \frac{1}{52} = \frac{4}{13}$$

or use the original definition of probability.

$$\frac{\text{number of aces and spades}}{\text{total number of cards}} = \frac{4+13-1}{52} = \frac{4}{13}$$

Multiplying Probabilities

$$\mathbf{P}(A \text{ and } B) = \mathbf{P}(A) \times \mathbf{P}(B)$$

provided A is not affected by the outcome of B and B is not affected by the outcome of A, i.e. A and B must be *independent*.

We shuffle two packs of cards and then draw a card from each of them. What is the probability that two aces are drawn?

$$\mathbf{P}(\text{ ace }) = \frac{1}{13}$$

$$\mathbf{P}(\text{ ace and ace }) = \frac{1}{13} \times \frac{1}{13} = \frac{1}{169}$$

Non-independent Events

I have a single pack of cards. I draw a card, then draw a second card without putting the first card back in the pack. What is the probability that I draw two aces?

This time the probability that I get an ace as the second card is affected by whether or not I removed an ace from the pack when I drew the first card.

We use the notation $\mathbf{P}(B|A)$ to denote the probability that B happens, given that we know that A happened. This is called a **conditional probability**.

$$\mathbf{P}(A \text{ and } B) = \mathbf{P}(A|B)\,\mathbf{P}(B)$$

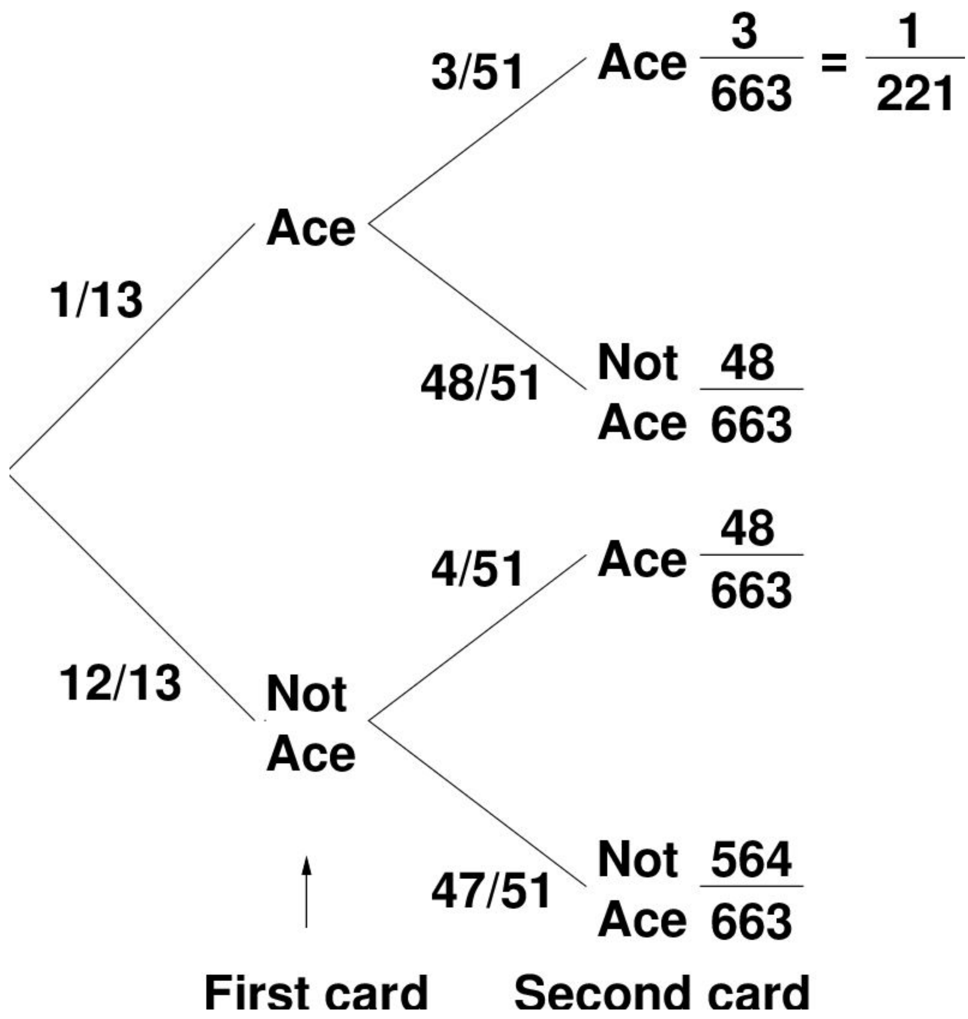
Thus:

$$\mathbf{P}(\text{ [second} = \text{ace] and [first} = \text{ace] })$$

$$= \mathbf{P}(\text{ second} = \text{ace} \mid \text{first} = \text{ace })\,\mathbf{P}(\text{ first} = \text{ace })$$

Tree Diagrams

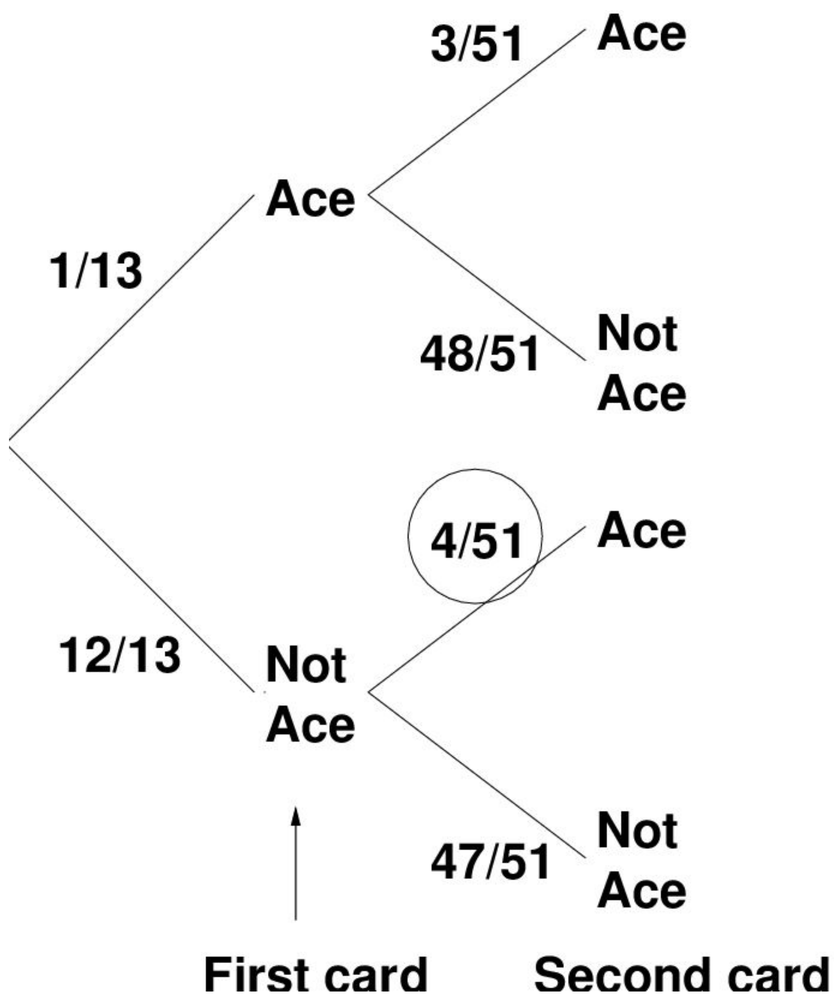
We can use a tree diagram to help us work this out.



The probability that both cards are aces $= \frac{1}{221}$.

Another Example

I have a single pack of cards. I draw a card, then draw a second card without putting the first card back in the pack. What is the probability that the second card is an ace, given that the first card was not an ace?



$$\mathbf{P}(\text{ second} = \text{ace} \mid \text{first} = \text{not ace }) = \frac{4}{51}$$

Notation

Intersection	$\bigcap$	AND
Union	$\bigcup$	OR

Thus the conditional probability formula

$$\mathbf{P}(A \text{ and } B) = \mathbf{P}(A|B)\,\mathbf{P}(B)$$

is more normally written

$$\mathbf{P}(A \cap B) = \mathbf{P}(A|B)\,\mathbf{P}(B)$$

and instead of

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

we write

$$\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B)$$

Ordering Objects

The number of different orders in which n unique objects can be placed is $n!$ (n factorial)

I have three cards with values 2, 3 and 4. They are shuffled into a random order. What is the probability they are in the order 2, 3, 4?

The number of possible orders for three cards is 3! The probability the cards are found in one specific order is therefore $\frac{1}{3!} = \frac{1}{6}$.

Permutations

$$_nP_r = \frac{n!}{(n-r)!}$$

is the number of ways of choosing r items from n when the order of the chosen items matters.

Ten people are involved in a race. I wish to make a poster for every possible winning combination of gold, silver and bronze medal winners. How many posters will I need?

We need to know the number of ways of choosing three people out of 10, taking account of the order. This is

$$_{10}P_3 = \frac{10!}{7!} = 10 \times 9 \times 8 = 720$$

So I would need rather a lot of posters.

Combinations

$$_nC_r = \frac{n!}{(n-r)!r!}$$

is the number of ways of choosing r items from n when the order of the chosen items does not matter.

I have a single pack of cards. I draw a card, then draw a second card without putting the first card back in the pack. What is the probability that I draw two aces?

The number of ways of drawing 2 cards from 52 is $_{52}C_2$.

The number of ways of getting two aces is the number of ways of drawing 2 aces from the 4 aces in the pack. This is $_4C_2$.

The probability that I draw two aces is therefore

$$\frac{\text{num ace pairs}}{\text{num pairs}} = \frac{_4C_2}{_{52}C_2} = \frac{4!}{2!\,2!} \times \frac{50!\,2!}{52!}$$
$$= \frac{4 \times 3}{52 \times 51} = \frac{1}{221}$$

Lottery Example 1

What is the probability of winning the jackpot in the national lottery? There are 49 balls and you have to match all six to win.

Method 1:

$$\frac{6}{49} \times \frac{5}{48} \times \frac{4}{47} \times \frac{3}{46} \times \frac{2}{45} \times \frac{1}{44} = \frac{1}{13983816}$$

Method 2:

$$\frac{1}{\text{number of ways of choosing 6 balls from 49, where order does not matter}} = \frac{1}{{}_{49}C_6}$$

$$\frac{1}{{}_{49}C_6} = \frac{6!\,43!}{49!} = \frac{6 \times 5 \times 4 \times 3 \times 2 \times 1}{49 \times 48 \times 47 \times 46 \times 45 \times 44}$$
$$= \frac{1}{13983816}$$

Lottery Example 2

What is the probability of winning £10 by matching exactly 3 balls in the national lottery

Method 1:

Work out the probability of matching them in a particular order: the first 3 balls that are drawn win, the remaining 3 do not.

$$\frac{6}{49} \times \frac{5}{48} \times \frac{4}{47} \times \frac{43}{46} \times \frac{42}{45} \times \frac{41}{44}$$

Then multiply this by the number of possible ways of picking the 3 winning balls among the 6 balls that are drawn.

$$\left.\begin{array}{llllll}
\checkmark & \checkmark & \checkmark & \boldsymbol{\times} & \boldsymbol{\times} & \boldsymbol{\times} \\
\checkmark & \checkmark & \boldsymbol{\times} & \checkmark & \boldsymbol{\times} & \boldsymbol{\times} \\
\checkmark & \checkmark & \boldsymbol{\times} & \boldsymbol{\times} & \checkmark & \boldsymbol{\times} \\
\checkmark & \checkmark & \boldsymbol{\times} & \boldsymbol{\times} & \boldsymbol{\times} & \checkmark \\
\checkmark & \boldsymbol{\times} & \checkmark & \checkmark & \boldsymbol{\times} & \boldsymbol{\times} \\
\multicolumn{6}{l}{\ldots \text{ etc.}}
\end{array}\right\} {}_6C_3 = 20$$

Hence: $\frac{6}{49} \times \frac{5}{48} \times \frac{4}{47} \times \frac{43}{46} \times \frac{42}{45} \times \frac{41}{44} \times {}_6C_3 = \frac{1}{56.7}$

Matching only 3 balls, method 2:

$$\frac{\text{number of ways we can win}}{\text{total possible number of outcomes}}$$

Thinking about all the balls in the lottery machine, we consider:

$$\frac{\begin{pmatrix}\text{The number of ways the} \\ \text{lottery machine can pick} \\ \text{3 balls matching some of} \\ \text{the 6 numbers on our} \\ \text{ticket.}\end{pmatrix}\begin{pmatrix}\text{The number of ways the} \\ \text{lottery machine can pick} \\ \text{3 balls from the 43 balls} \\ \text{not on our ticket.}\end{pmatrix}}{\begin{pmatrix}\text{The total number of ways of picking 6} \\ \text{balls out of the 49 in the machine.}\end{pmatrix}}$$

$$\begin{aligned}
&= \frac{{}_6C_3\,{}_{43}C_3}{{}_{49}C_6} = \frac{6!\,43!\,43!\,6!}{49!\,40!\,3!\,3!\,3!} \\
&= \frac{43 \times 42 \times 41 \times 6 \times 5 \times 4 \times 6 \times 5 \times 4}{49 \times 48 \times 47 \times 46 \times 45 \times 44 \times 3 \times 2} \\
&= \frac{1}{56.7}
\end{aligned}$$

Section 9: Summary

$\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B)$ if A and B are mutually exclusive outcomes.

$\mathbf{P}(A \cap B) = \mathbf{P}(A) \times \mathbf{P}(B)$ provided A and B are independent.

$\mathbf{P}(A \cap B) = \mathbf{P}(A|B)\,\mathbf{P}(B)$

$\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B)$

The number of different orders in which n unique objects can be placed is $n!$

Permutations: ${}_nP_r = \frac{n!}{(n-r)!}$ is the number of ways of choosing r items from n when the order of the chosen items matters.

Combinations: ${}_nC_r = \frac{n!}{(n-r)!r!}$ is the number of ways of choosing r items from n when the order of the chosen items does not matter.

Section 10

Statistics

In this section we summarise the key issues in pages 14–20 of the basic probability teach-yourself document. This presentation is intended to be reinforced by the examples in the teach-yourself document and questions 13 and 14 in examples paper 9.

The main focus is on the mean and standard deviation of a probability distribution. We also explain how to calculate a range within which we are (say) 95% sure that the true value of an experimental reading will lie.

Mean

The mean μ, of a population of values x_i (where i goes from 1 to N), is defined as

$$\mu = \frac{\text{Sum of all the values}}{\text{Number of values}} = \frac{\sum_{i=1}^{N} x_i}{N}$$

Imagine a pack of cards with all the jokers and picture cards removed. We are only concerned with the numerical value of the cards. We have four each of all the numbers from one to ten so $N = 40$.

$$\text{Arithmetic mean: } \mu = \frac{\sum_{i=1}^{N} x_i}{N} = \frac{220}{40} = 5.5$$

The mean is a measure of the *central tendency* or *location* of the population.

Variance and Standard Deviation

Variance and standard deviation are measures of the *spread* of the distribution. The variance is the average squared difference between each value and the mean. The population variance is usually given the symbol σ^2.

$$\text{Variance: } \sigma^2 = \frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N}$$

The standard deviation (SD) is the square root of the variance. The population standard deviation is usually given the symbol σ.

$$\text{Standard Deviation: } \sigma = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N}}$$

We can work out the variance and standard deviation of the values on our set of forty cards.

$$\sigma^2 = \frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N} = \frac{330}{40} = 8.25$$

$$\sigma = \sqrt{\frac{330}{40}} = 2.8723$$

Sample

Until now we have assumed that we can see all the cards at once. Now we are going to change the game. Imagine that someone else is holding the cards and allowing us to pick one at random, note its value and then replace it. Using this pick-and-replace process we can view a *sample* of the cards. This sample can be of any size as the cards are picked at random and replaced. Assume that the sample size is n.

The challenge is to estimate the mean and standard deviation of the original numbers on the cards based only on what we see in the sample. Here are the formulae that enable us to do this.

Estimate of Mean (based on sample):

$$m = \frac{\sum_{i=1}^{n} x_i}{n}$$

Estimate of Standard Deviation (based on sample):

$$s = \sqrt{\frac{\sum_{i=1}^{n} (x_i - m)^2}{n-1}}$$

Note: the appearance of $(n-1)$ in the denominator rather than n.

Convenient Formula for s

In general, s, the standard deviation of the underlying population estimated from a sample is called the *"sample standard deviation"*.

There is a convenient formula for calculating s.

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - m)^2}{n-1}}$$

$$= \sqrt{\frac{\left(\sum_{i=1}^{n} x_i^2\right) - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2}{n-1}}$$

Standard Deviation from a Sample

Ten cards are selected individually from our special reduced pack of 40 cards (described on the *Mean* page), noted and replaced in the pack. This gives a sample size $n = 10$. The values of the cards are:

$$\begin{array}{rrrrr} 10 & 3 & 4 & 3 & 5 \\ 4 & 1 & 5 & 8 & 5 \end{array}$$

We wish to estimate the mean m, and standard deviation s, of the values on all the cards, based only on knowledge of this sample.

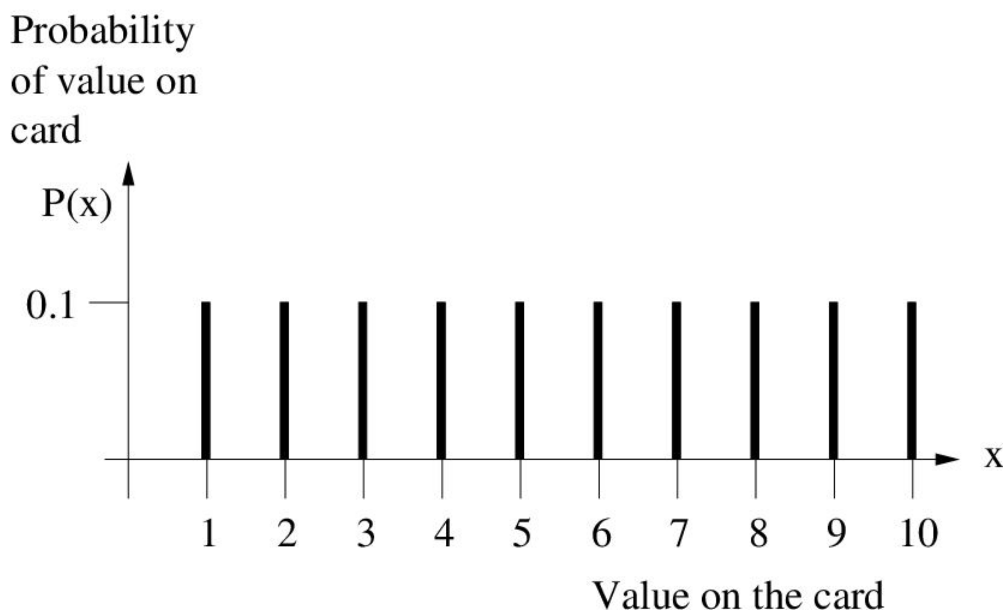
$$m = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{48}{10} = 4.8$$

$$s = \sqrt{\frac{\left(\sum_{i=1}^{n} x_i^2\right) - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2}{n-1}}$$

$$= \sqrt{\frac{(290) - \frac{1}{10}(48)^2}{9}} = 2.5734$$

Discrete Probability Distributions

Consider picking a card (again from the pack described on the *Mean* page), noting its value and then replacing it in the pack. We can compute the probability of picking each of the possible values.



This is a probability distribution. In this case it is a *discrete* distribution because the cards can only carry certain integer values. Notice that the sum of all the histogram bars is $10 \times 0.1 = 1$. There are ten possible outcomes and they each have a probability of $1/10$. This is called a *uniform* distribution.

Mean and SD from the Distribution

The probability distribution is a property of the population of the numbers on the cards. Knowing the complete probability distribution enables us to calculate the mean μ and the standard deviation σ exactly.

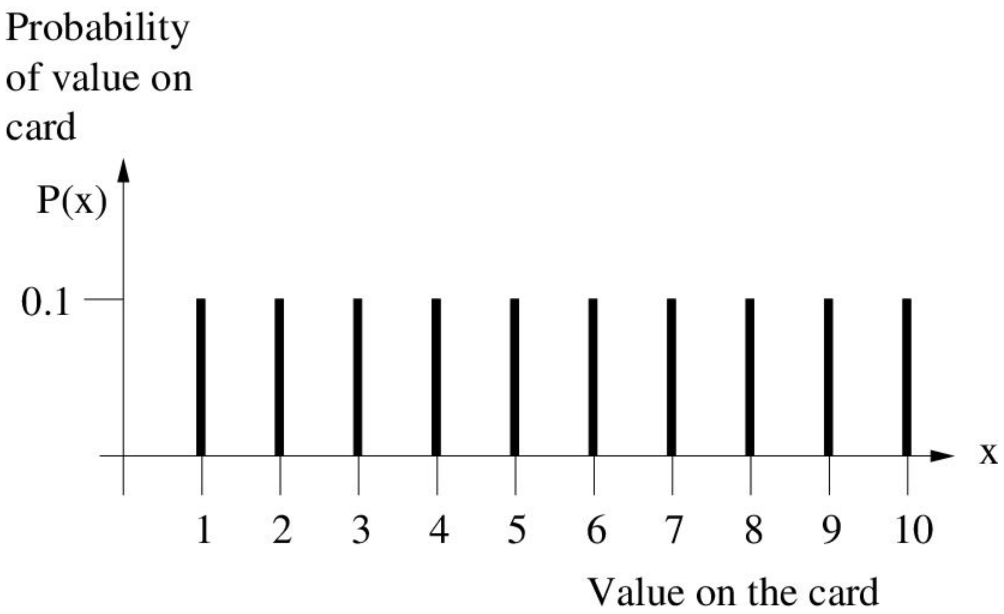
Let x_j represent each of the *different* values that are printed on the cards and M equal the number of these different values. In our example $M = 10$. If $P(x_j)$ is the probability of x_j occurring, e.g. $P(1) = 0.1$ etc., then

$$\text{Arithmetic mean: } \mu = \sum_{j=1}^{M} x_j P(x_j)$$

$$\text{Variance: } \sigma^2 = \sum_{j=1}^{M} (x_j - \mu)^2 P(x_j)$$

$$\text{Standard Deviation: } \sigma = \sqrt{\sum_{j=1}^{M} (x_j - \mu)^2 P(x_j)}$$

Example



We can see from the histogram that $P(x_j) = 0.1$ for all the values on the cards (i.e. for all j). In this particular case, the values x_j are the same numerically as the index j, so we can substitute $x_j = j$. Hence

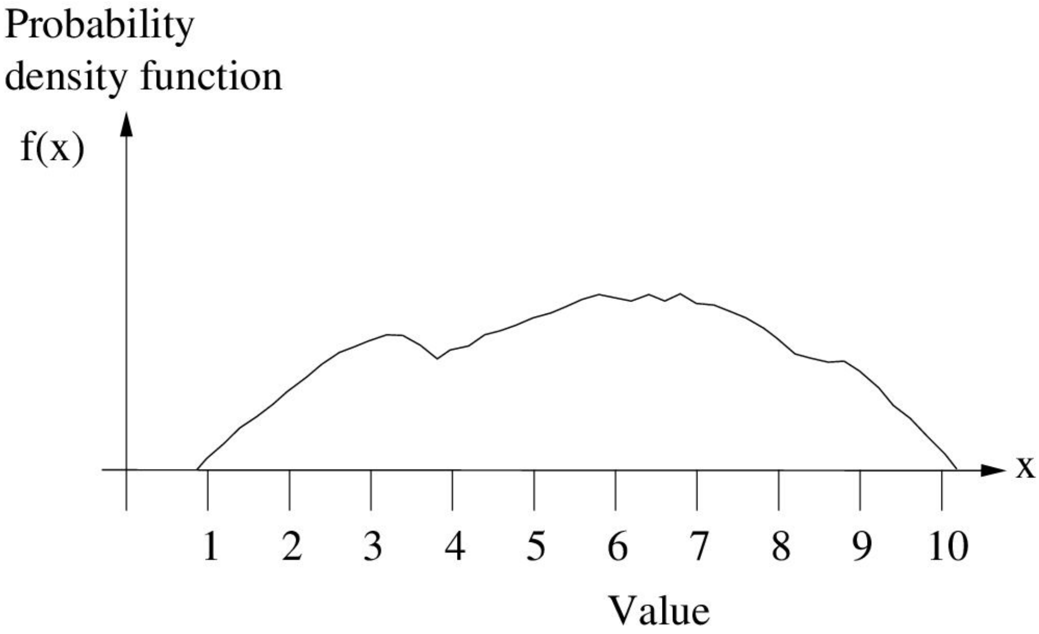
$$\mu = \sum_{j=1}^{M} x_j P(x_j) = \sum_{j=1}^{10} j \times 0.1 = 5.5$$

Now we use this value of μ in the formula for standard deviation.

$$\begin{aligned}\sigma &= \sqrt{\sum_{j=1}^{M} (x_j - \mu)^2 P(x_j)} \\ &= \sqrt{\sum_{j=1}^{10} (j - 5.5)^2 \times 0.1} = 2.8723\end{aligned}$$

Continuous Probability Distributions

If you have an outcome that can take any real value (rather than a finite number of discrete values) this can be described by a probability density function (PDF).



Here the total area under the curve must be 1 and the probability of x taking a value in the range from (say) 6 to 7 is given by the integral (i.e. area) between 6 and 7. More generally:

$$\text{The probability of } (a < x < b) = \int_a^b f(x)\,dx$$

Examples of continuous random variables: the weight of a sample, the time for a physical process to complete, an output voltage.

Mean and SD from PDF

It is also possible to calculate the mean (μ) and standard deviation (σ) of a distribution from its probability density function.

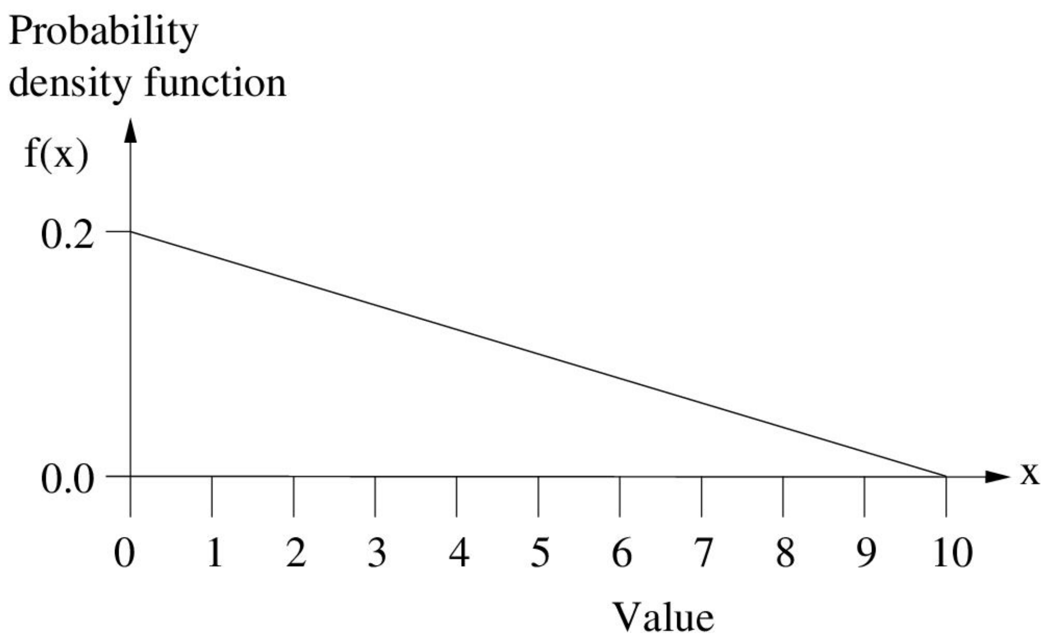
$$\mu = \int_{-\infty}^{+\infty} x f(x)\, dx$$

$$\sigma = \sqrt{\int_{-\infty}^{+\infty} (x - \mu)^2 f(x)\, dx}$$

Knowing the probability density function enables us to calculate the mean μ and the standard deviation σ exactly.

Example using a PDF

Consider a machine than makes widgets which are supposed to be a particular length. Unfortunately, the machine often makes widgets that are slightly too long; it never makes widgets that are too short. The graph below shows the probability density function for the number of millimetres that a widget is too long.



Widget Example 1

1. What is the probability that a widget is less than 1 mm too long?

For $0 \leq x \leq 10$ we can see that $f(x) = 0.2 - 0.02x$, hence:

$$\begin{aligned} P(0 < x < 1) &= \int_0^1 f(x)\,dx \\ &= \int_0^1 (0.2 - 0.02x)\,dx = 0.19 \end{aligned}$$

So the probability of a widget being less than 1 mm too long is 0.19.

2. Calculate the mean and standard deviation of the distribution of excess lengths.

$$\mu = \int_0^{10} x(0.2 - 0.02x)\,dx = 3.3333$$

$$\sigma = \sqrt{\int_0^{10} (x - 10/3)^2 (0.2 - 0.02x)\,dx} = 2.3570$$

Widget Example 2

3. What is the probability that a widget is produced with an excess length within one standard deviation from the mean excess length?

We wish to calculate the probability of an excess length in the range $3.3333 - 2.3570$ to $3.3333 + 2.3570$, which is given by:

$$\begin{aligned} \mathrm{P}(x \text{ within one } \sigma \text{ of } \mu) &= \int_{0.9763}^{5.6904} 0.2 - 0.02x \, dx \\ &= 0.6285 \end{aligned}$$

4. The manufacturer wants to quote an excess length that he is sure 95% of the widgets produced will be shorter than. What should it be?

We need to solve for d in:

$$\begin{aligned} 0.95 &= \int_0^d 0.2 - 0.02x \, dx \\ \Rightarrow 0.95 &= 0.2d - 0.01d^2 \end{aligned}$$

So $d = 10 - \sqrt{5} = 7.7639$ mm

Standard Deviation of a Sample Mean

Suppose x is a random variable with

$$\begin{aligned} \text{mean} &= \mu \\ \text{standard deviation} &= \sigma \end{aligned}$$

We have n samples of x: $\left[x_1, x_2, x_3, \ldots x_n\right]$.

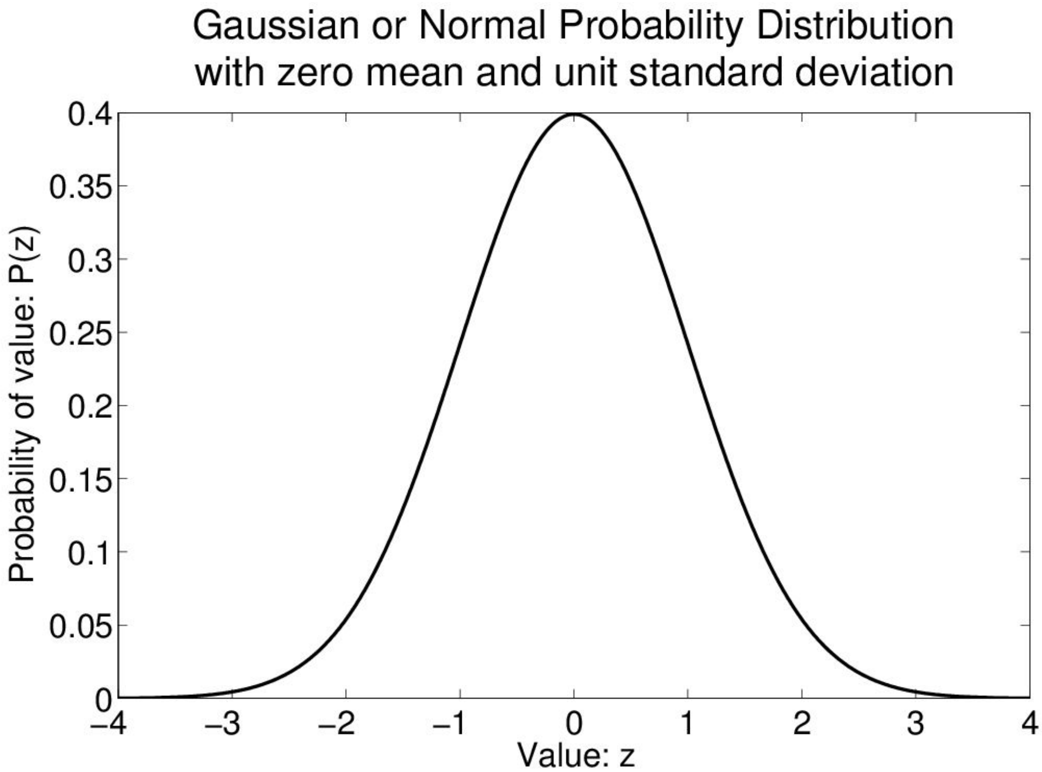
Let $\overline{x}$ be the average of these samples $= \sum_{i=1}^{n} \frac{x_i}{n}$.

Now the key point is that $\overline{x}$ is also a random variable.

The mean of $\overline{x}$ is μ.

The standard deviation of $\overline{x}$ is $\frac{\sigma}{\sqrt{n}}$.

Normal Probability Distribution



The Normal distribution is a symmetric distribution with two parameters: the *mean* μ and the *standard deviation* σ.

$$P(z) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(z-\mu)^2}{2\sigma^2}\right)$$

If you add together a sufficient number (say 30), of independent random variables that are identically distributed, the sum will conform to the Normal distribution. This is called the "*central limit theorem*".

Normal Distribution Example

Consider a laboratory experiment that results in a single real output x, each time that we perform it. In theory, it should produce the same output each time but in practice x varies slightly because of noise in the measurement system.

If we repeat the experiment at least 30 times and average the result ($\overline{x} = m = \sum_{i=1}^{n} x_i / n$) then we can say the following things about the way that $\overline{x}$ is distributed.

- Provided $n > 30$ it is reasonable to assume that $\overline{x}$ is Normally distributed.
- The standard deviation of $\overline{x}$ will be a factor of $\sqrt{n}$ less than the standard deviation of the original experimental data. Hence:

Estimate of standard deviation of $\overline{x}$:

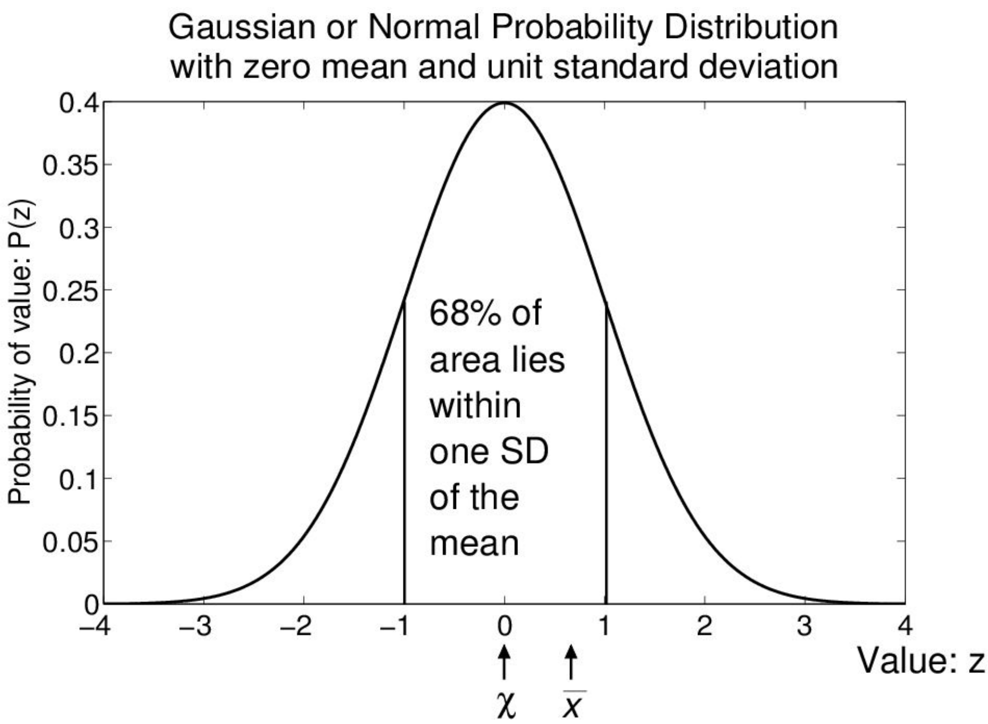
$$s_{(\overline{x})} = \frac{1}{\sqrt{n}} \times \sqrt{\frac{\left(\sum_{i=1}^{n} x_i^2\right) - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2}{n-1}}$$

Normal Distribution Example 2

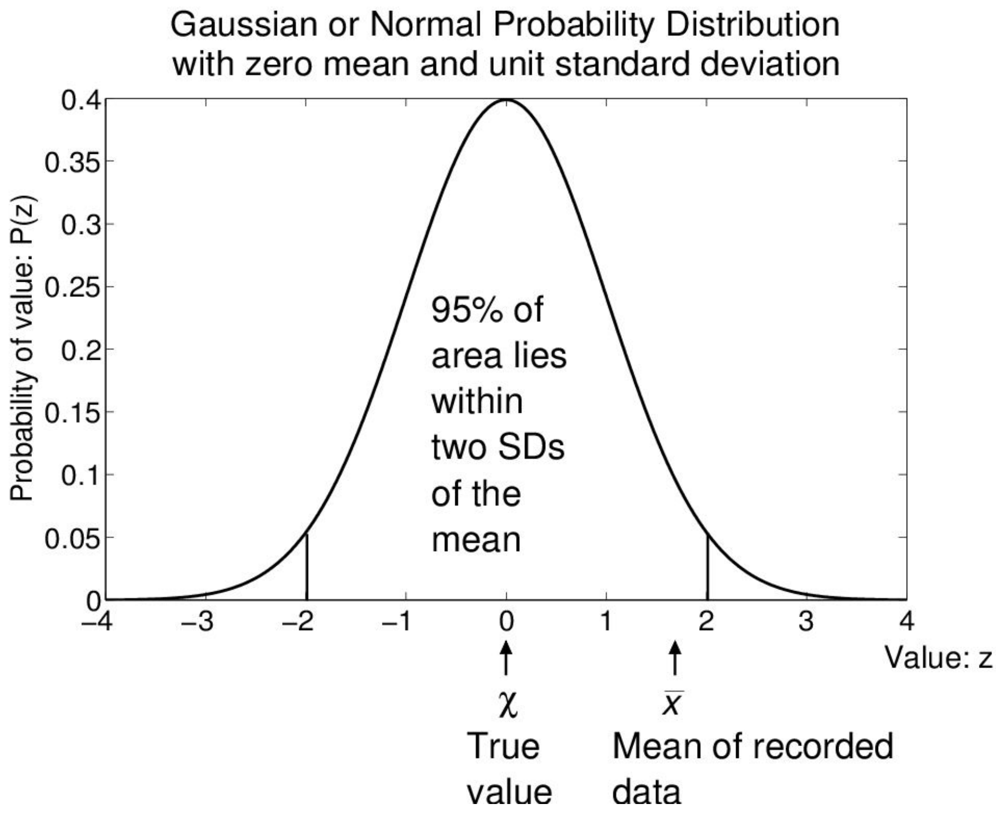
If it is fair to assume that the error in the original experimental data is *unbiased* ('unbiased' indicates that the expected value of the error is equal to the true mean of the error), then the standard deviation of $\overline{x}$ gives us useful information about the error it is likely to contain.

Let $s_{(\overline{x})}$ be our estimate of the standard deviation of $\overline{x}$ and let $\mathcal{X}$ be the (unknown) true value of the thing we are intending to measure.

50%	of the time, $\overline{x}$ will lie within	$0.67 s_{(\overline{x})}$	of $\mathcal{X}$
68%	of the time, $\overline{x}$ will lie within	$s_{(\overline{x})}$	of $\mathcal{X}$
95%	of the time, $\overline{x}$ will lie within	$2 s_{(\overline{x})}$	of $\mathcal{X}$
99.73%	of the time, $\overline{x}$ will lie within	$3 s_{(\overline{x})}$	of $\mathcal{X}$



Normal Distribution Example 3



If we repeat an experiment 30 times and

- the average result $\overline{x} = 3.0279$,
- we calculate an estimate of the standard deviation of the experimental error as 0.2036
- then our estimate of the standard deviation of $\overline{x}$ will be $0.2036/\sqrt{30} = 0.0372$.

The true experimental result will have a 95% chance of lying within two standard deviations from the mean, i.e. in the range from 2.9536 to 3.1023.

Section 10: Summary

PDF	probability density function
SD	standard deviation
μ	mean of a distribution (or from a PDF)
σ^2	variance of a distribution (or from a PDF)
σ	SD of a distribution (or from a PDF)
$m = \overline{x}$	estimate of μ based on a sample of x values
s	estimate of σ based on a sample of x values
$s_{(\overline{x})}$	estimate of SD of $\overline{x}$ based on sample of x
	$s_{(\overline{x})} = s/\sqrt{n}$ where n is the sample size.

If we are prepared to do an experiment at least 30 times and believe our results to be unbiased, we can use the "central limit theorem" together with the shape of the Normal distribution to calculate a range within which the true result is likely to lie. We can make a statement of the form: "There is a probability of p that the true result lies between x_{low} and x_{high}." This is called a confidence interval.

Acknowledgements

These notes are currently almost identical (but differ in parts) to Prof. Prager's (RWP) notes for this course (pre 2013). RWP's notes were inspired by Prof. Woodhouse's notes from an earlier version of the course.