Practice Questions

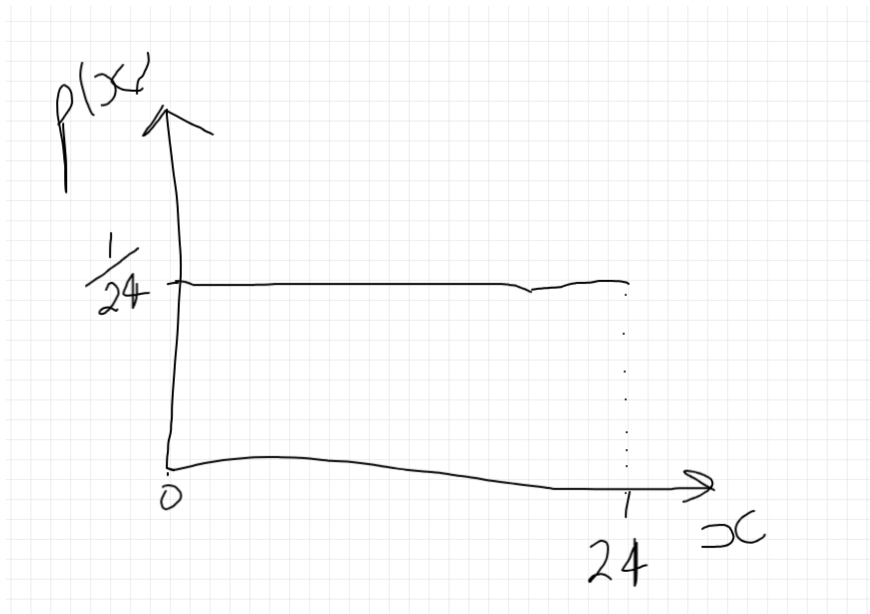
Thomas Lawson

November 2, 2025

1 Sheet 1

Question 1) X is a continuous random variable representing the time of day at which an event occurs, measured in hours 0 to 24. The event is equally likely to occur at any time.

(i) Sketch the probability density function, $p(x)$



(ii) Write out the equation for $p(x)$.

We know that $p(x)$ has a value of $\frac{1}{24}$ if and only if $0 \leq x \leq 24$, otherwise it has a value of 0 (no chance of it happening).

With this we can form the following equation:

$$p(x) = \begin{cases} \frac{1}{24} & \text{if } 1 \leq x \leq 24 \\ 0 & \text{otherwise} \end{cases}$$

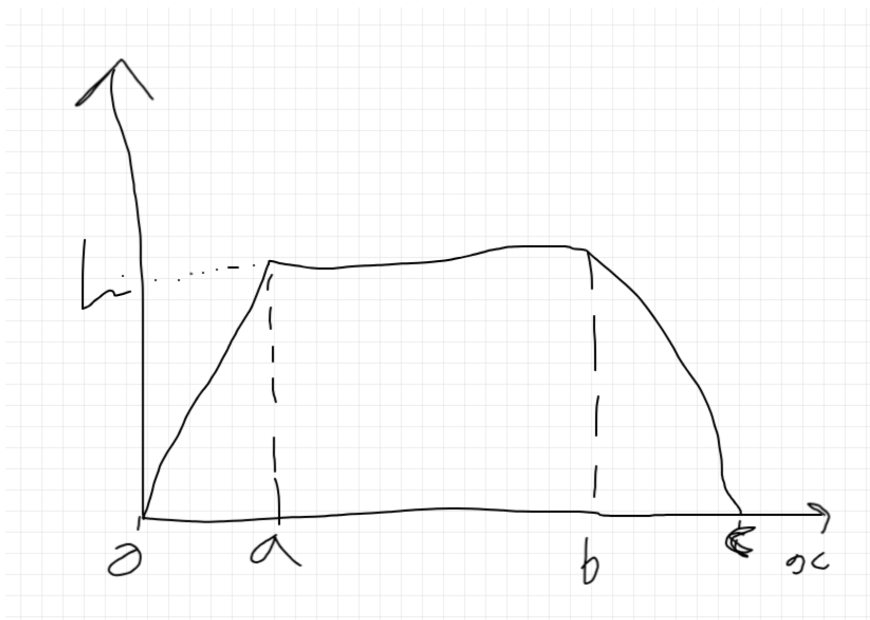
(iii) What is the probability that the event happening between 8:00 and 12:00?

This is a 4 hour period. A 4 hour period makes up $\frac{1}{6}$ hours of a day.

The event is equally likely to happen at any time of the day and the area under the PDF is 1 $\therefore p(8 \leq x \leq 12) = 1 \times \frac{1}{6} = \frac{1}{6}$

Question 2) A pdf, p(x), has a value that linearly increases from x = 0 to x = a, remains constant from x = a to x = b and then linearly decreases from x = b to x = c, where 0 < a < b < c. p(x) = 0 for x < 0 and for x > c. The pdf has no discontinuities.

(i) Sketch $p(x)$



(ii) What is the value of p(x) for x = (a + b)/2

The equation we've been give here is almost that of the area of a trapezium ($A = \frac{a-b}{2} \times h$) except there is no h. If you divide both sides by h then that gives us

$$\frac{A}{h} = \frac{a+b}{2}$$

The RHS of this equation is the same as the RHS of the equation of given in the question which means that $x = \frac{A}{h}$

A reprents the area of the trapezium, the pdf has the shape of the trapezium and all pdfs have an area under the curve equal to one $\therefore$ $A = 1$.

We can now rearrange to get h

$$\frac{1}{h} = \frac{a+b}{2}$$
$$1 = \frac{a+b}{2} \times h$$
$$\frac{1}{\frac{a+b}{2}} = h$$
$$\frac{2}{a+b} = h$$

I can now substitute the values corresponding to the sketch, not the original equation of the trapezium giving us:

$$\frac{2}{b-a+c} = h$$

(iii) What is the probability that x has a value less than a

We can get this by doing $h \times a \times 0.5$ just like when calculating the area of a triangle.

$$h \times a \times \frac{1}{2} = \frac{2}{b-a+c} \times a \times \frac{1}{2} = \frac{1}{b-a+c} \times a = \frac{a}{b-a+c}$$

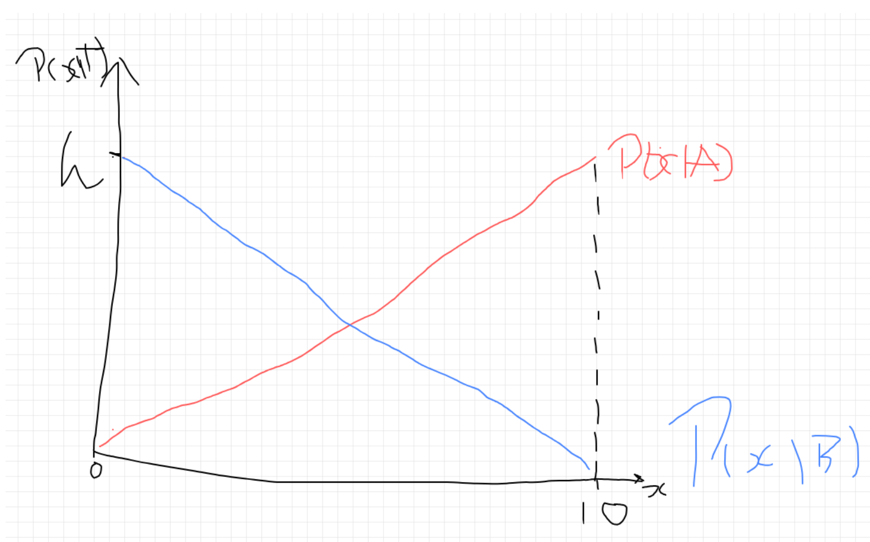
(iv) What is the probability that x has a value greater than b

We can get this by by doing $h \times (c-b) \times 0.5$ just like when calculating the area of a triangle.

$$h \times (c-b) \times 0.5 = \frac{2}{b-a+c} \times (c-b) \times \frac{1}{2} = \frac{c-b}{b-a+c}$$

Question 3) On an alien planet there are two types of tree, type A and type B. The heights of type A have a pdf that linearly decreases from a maximum at a height of 0 m down to a value of 0 at a height of 10 m. The heights of type B have a pdf that is 0 at 0 m and linearly increases to a maximum value at 10 m. There are no trees taller than 10 m. Assume also that there are an equal number of A and Bs on the planet.

(i) Sketch the pdfs of the heights for type A and type B, ie., p(x|A) and p(x|B)



(ii) What is the probability of a type A tree being between 3 and 4 meters tall?

We're trying to find the area under the graph for $p(3 \leq x \leq 4|\,A)$.

This area forms a trapezium, which means we can solve the area using $A = \frac{b+a}{2} \times h_t$. We can easily calculate h_t by doing $4-3$ which gives us 1.

To find the values of $p(x=4\,|\,A)$ and $p(x=3\,|\,A)$ we can start by finding the equation for $p(x\,|\,A)$ which is $\frac{1}{5} \times \frac{10-x}{10}$.

Using this we can get the values for $p(x=4\,|\,A)$ and $p(x=3\,|\,A)$.

$$p(x=4\,|\,A) = \frac{1}{5} \times \frac{10-4}{10} = \frac{1}{5} \times \frac{6}{10} = \frac{6}{50}$$
$$p(x=3\,|\,A) = \frac{1}{5} \times \frac{10-3}{10} = \frac{1}{5} \times \frac{7}{10} = \frac{7}{50}$$

Now we can find the probability of tree type A being between 3 and 4

$$\frac{\frac{7}{50}+\frac{6}{50}}{2}\times 1=\frac{\frac{13}{50}}{2}=\frac{13}{100}$$

(iii) A tree is found and observed to be 6 meters tall. What is the probability that it is of type A?

We can deduce that the equation for $p(x|B)$ is $\frac{1}{5}\times\frac{x}{10}$.

Since there are the same number of type A and type B trees, we can divide the probability of $p(x=6|A)$ by $p(x=6|A)+p(x=6|B)$ to get our answer.

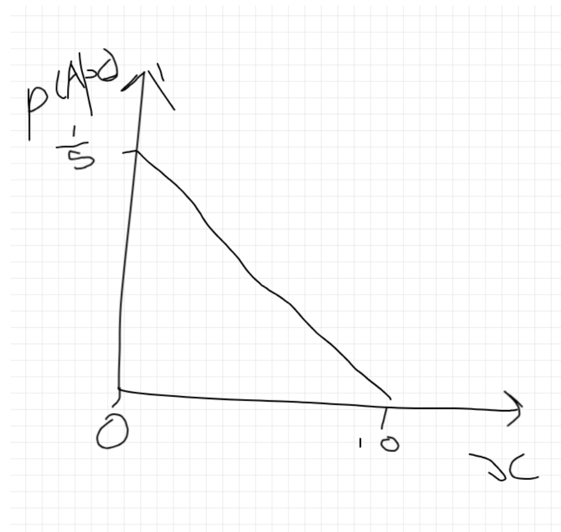
$$\frac{p(x=6|A)}{p(x=6|A)+p(x=6|B)}=\frac{\frac{1}{5}\times\frac{10-6}{10}}{\frac{1}{5}\times\frac{10-6}{10}+\frac{1}{5}\times\frac{6}{10}}=\frac{\frac{10-6}{10}}{\frac{10-6}{10}+\frac{6}{10}}=\frac{\frac{4}{10}}{\frac{4}{10}+\frac{6}{10}}=\frac{4}{4+6}=\frac{4}{10}$$

(iv) Sketch the posterior probability p(A|x) as a function of x.

Pre-Sketch Notes

$p(A|x)$ must have a value of 0.5 when $x=0$, as at that point there should be no type B trees. It should be 0 when $x=10$. The line should start at 0 and end at 10.

The sketch should look identical to the one done of $p(x|A)$ earlier in the question.



(v) Using Bayes' decision rule, what would be the threshold height that a classifier would use to seperate Type A and Type B?

I'm looking to find a value for x where we start getting $p(A|x)<p(B|x)$ instead of $p(A|x)>p(B|x)$.

We haven't calculated $p(A|x)$ nor $p(B|x)$ yet but we can tell that they hold that same values as $p(x|A) and p(x|B)$ respectively

Here are $p(A|x)$ and $p(B|x)$ represented by Baye's rule

$$p(A|x)=\frac{p(x|A)\times p(A)}{P(x)}$$

$$p(B|x)=\frac{p(x|B)\times p(B)}{P(x)}$$

The question is asking me to calculate the threshold height using Bayes' decision rule so I'll set the two as equal to each other and to obtain the value of x where they return the same value which should represent the threshold height.

$$\frac{p(x|A)\times p(A)}{P(x)}=\frac{p(x|B)\times p(B)}{P(x)}$$

Cancel out $p(x)$

$$p(x|A) \times p(A) = p(x|B) \times p(B)$$

Cancel out $p(A)$ and $p(B)$ as there are the same amount of each $\therefore$ they have the same value.

$$p(x|A) = p(x|B)$$

Substitute in our equations from earlier

$$\frac{1}{5} \times \frac{10-x}{10} = \frac{1}{5} \times \frac{x}{10}$$

Now we can just simplify

$$\frac{10-x}{10} = \frac{x}{10}$$
$$\frac{10-x}{10} = \frac{x}{10}$$
$$10 - x = x$$
$$10 = 2x$$
$$5 = x$$

Now we know that the threshold height is 5, if $x < 5$ it can be classied as type A, otherwise as B.

(vi) What would be the probability of a classification error for this classifier?

First attempt (Wrong)

I'll calculate this by getting the area under each of the two PDFs on each side of the threshold height. I will then use to to calculate probability of each side being incorrectly labelled.

I'll start by calculating area under $p(A|x \leq 5)$

$$\frac{\frac{1}{5} + \frac{1}{5} \times \frac{10-5}{10}}{2} \times 5 = \frac{\frac{1}{5} + \frac{1}{5} \times \frac{1}{2}}{2} \times 5 = \frac{\frac{1}{5} + \frac{1}{10}}{2} \times 5 = \frac{1 + \frac{1}{5}}{2} = \frac{\frac{6}{5}}{2} = \frac{6}{10}$$

This means that $p(A|x \geq 5) = 0.4$

The graph mirrors at $x = 5$ so we can set inverse values for B, $p(B|x \leq 5) = 0.4$ and $p(B|x \leq 5) = 0.6$
Now we can calculate how often the classifier will be wrong.

$$(\frac{0.4}{0.4 + 0.6} \times 0.5) \times 2 = \frac{4}{10}$$

Second attempt (This one is right)

I'll calculate this by getting the area under each of the two PDFs on each side of the threshold height. I will then use to to calculate probability of each side being incorrectly labelled.

I'll start by calculating area under $p(A|x \leq 5)$

$$\frac{\frac{1}{5} + \frac{1}{5} \times \frac{10-5}{10}}{2} \times 5 = \frac{\frac{1}{5} + \frac{1}{5} \times \frac{1}{2}}{2} \times 5 = \frac{\frac{1}{5} + \frac{1}{10}}{2} \times 5 = \frac{\frac{3}{10}}{2} \times 5 = \frac{3}{20} \times 5 = \frac{15}{20} = \frac{3}{4}$$

This means that $p(A|x \geq 5) = \frac{1}{4}$

The graph mirrors at $x = 5$ so we can set inverse values for B, $p(B|x \leq 5) = \frac{1}{4}$ and $p(B|x \leq 5) = \frac{3}{4}$

Now we can calculate how often the classifier will be wrong.

$$(\frac{\frac{1}{4}}{\frac{1}{4} + \frac{3}{4}} \times 0.5) \times 2 = \frac{1}{4}$$

(vii) If it is now discovered that there are in fact twice as many type B trees as type A trees, i.e. the prior probabilities are not equal. What would be the new best decision threshold?

To get the new value for the decision threshold, we can just continue from midway part v of this question at the point where the prior probabilities were cancelled out except we don't cancel out for this questions as $p(A)$ and $p(B)$ no longer hold the same value.

$$p(x|A) \times p(A) = p(x|B) \times p(B)$$

We can now substitue in

$$\frac{1}{5} \times \frac{10-x}{10} \times \frac{1}{3} = \frac{1}{5} \times \frac{x}{10} \times \frac{2}{3}$$

Multiply both sides by 15

$$\frac{10-x}{10} = \frac{2x}{10}$$

Multiply both sides by 10

$$10 - x = 2x$$

Add x to both sides

$$10 = 3x$$

We get a threshold value for x by dividing both sides by 3

$$\frac{10}{3} = x$$

(viii) What would be the probability of classification error in this unequal prior scenario?

With this new decision threshold, the classifier should classify any tree below $\frac{10}{3}$ meters as tree type A, otherwise it will class it as type B.

I'll start by calculating the probability of $p(x \leq \frac{10}{3}|A)$

$$p(x \leq \frac{10}{3}|A) = \frac{\frac{1}{5} + \frac{1}{5} \times \frac{10-\frac{10}{3}}{10}}{2} \times \frac{10}{3} = \frac{\frac{10}{5} + \frac{1}{5} \times (10 - \frac{10}{3})}{6} = \frac{\frac{10}{5} + \frac{10}{5} - \frac{10}{15}}{6} = \frac{\frac{60}{15} - \frac{10}{15}}{6} = \frac{\frac{50}{15}}{6} = \frac{50}{90} = \frac{5}{9}$$

From this we can deduce that $p(x \geq \frac{10}{3}|A) = \frac{4}{9}$

Now we calculate the probability of $p(x \leq \frac{10}{3}|B)$

$$p(x \leq \frac{10}{3}|B) = \frac{0 + \frac{1}{5} \times \frac{\frac{10}{3}}{10}}{2} \times \frac{10}{3} = \frac{\frac{1}{5} \times \frac{1}{3}}{2} \times \frac{10}{3} = \frac{\frac{1}{15}}{2} \times \frac{10}{3} = \frac{\frac{10}{15}}{6} = \frac{\frac{2}{3}}{6} = \frac{2}{18} = \frac{1}{9}$$

From this we can deduce that $p(x \geq \frac{10}{3}|B) = \frac{8}{9}$

Now we can find the likelihood of $x \leq \frac{10}{3}$ being misslabeled.

$$\frac{\frac{1}{9} \times 2}{\frac{5}{9} + \frac{1}{9} \times 2} = \frac{\frac{2}{9}}{\frac{5}{9} + \frac{2}{9}} = \frac{2}{7}$$

Now we can find the likelihood of $x \geq \frac{10}{3}$ being misslabeled.

$$\frac{\frac{4}{9}}{\frac{4}{9} + \frac{8}{9} \times 2} = \frac{\frac{4}{9}}{\frac{4}{9} + \frac{16}{9}} = \frac{4}{20} = \frac{1}{5}$$

Now I will find the proportion of trees smaller than $\frac{10}{3}$

$$p(x \leq \frac{10}{3}) = \frac{\frac{5}{9} + \frac{1}{9} \times 2}{\frac{5}{9} + \frac{1}{9} \times 2 + \frac{4}{9} + \frac{8}{9} \times 2} = \frac{\frac{7}{9}}{\frac{7}{9} + \frac{20}{9}} = \frac{7}{27}$$

We can deduce now that $p(x \leq \frac{10}{3}) = \frac{20}{27}$
Now we can finally calculate the missclassification probability

$$p(x_{\text{missclassified}}) =$$

$$\begin{aligned}
p(x_{misslabeled}) &= p(x_{misslabeled}|x \leq \frac{10}{3}) \times p(x \leq \frac{10}{3}) + p(x_{misslabeled}|x \geq \frac{10}{3}) \times p(x \geq \frac{10}{3}) \\
&= \frac{2}{7} \times \frac{7}{27} + \frac{1}{5} \times \frac{20}{27} \\
&= \frac{14}{189} + \frac{20}{135} \\
&= \frac{14}{189} + \frac{20}{135} \\
&= \frac{210}{845} \\
&= \frac{2}{9}
\end{aligned}$$

(ix) What is the average height of the type A trees? Of the type B trees?

We will use the following equation to calculate the average heights (I had to google this):

$$\langle X \rangle = \int_a^b x\, p(x)\, dx$$

First we solve the average height for type A trees

$$
\begin{aligned}
\int_0^{10} x \times (\frac{1}{5} \times \frac{10-x}{10})\, dx &= \int_0^{10} x \times (\frac{10-x}{50})\, dx \\
&= \int_0^{10} \frac{x(10-x)}{50}\, dx \\
&= \int_0^{10} \frac{10x - x^2)}{50}\, dx \\
&= \int_0^{10} \frac{10x}{50} - \frac{x^2}{50}\, dx \\
&= \frac{1}{50}\int_0^{10} 10x - x^2\, dx \\
&= \frac{1}{50}[5x^2 - \frac{x^3}{3}]_0^{10} \\
&= \frac{1}{50}(5(10)^2 - \frac{10^3}{3}) \\
&= \frac{1}{50}(5 \times 100 - \frac{1000}{3}) \\
&= \frac{1}{50}(500 - \frac{1000}{3}) \\
&= \frac{1}{50}(\frac{500}{3}) \\
&= \frac{500}{150} \\
&= \frac{50}{15} \\
&= \frac{10}{3}
\end{aligned}
$$

So the average height for type A trees is $\frac{10}{3}$

Now we can use the same technique to calculate the average height for type B trees

$$
\begin{aligned}
\int_0^{10} x \times (\frac{1}{5} \times \frac{x}{10})\, dx &= \int_0^{10} \frac{x}{5} \times \frac{x}{10}\, dx \\
&= \int_0^{10} \frac{x^2}{50}\, dx \\
&= \frac{1}{50} \int_0^{10} x^2\, dx \\
&= \frac{1}{50} [\frac{x^3}{3}]_0^{10} \\
&= \frac{1}{50} \times \frac{10^3}{3} \\
&= \frac{10^3}{150} \\
&= \frac{1000}{150} \\
&= \frac{100}{15} \\
&= \frac{20}{3}
\end{aligned}
$$

Sheet 2

Question 1 - Vector operations

Consider the vectors $\vec{x}_1$ and $\vec{x}_2$ given below,

$$\vec{x}_1 = (3, 2, 1)^T, \quad \vec{x}_2 = (-2, 1, 0)^T$$

For each part below, first work out the answers by hand and then check your answers using Python and numpy.

(i) The lengths of $\vec{x}_1$ and $\vec{x}_2$.

We can find the magnitude of these two vectors by taking the square root of the sum of the squares of each element.

$$|\vec{x}_1| = \sqrt{3^2 + 2^2 + 1^2} = \sqrt{9 + 4 + 1} = \sqrt{14}$$
$$|\vec{x}_2| = \sqrt{(-2)^2 + 1^2 + 0^2} = \sqrt{4 + 1 + 0} = \sqrt{5}$$

(ii) The ℓ_1-norm distance from $\vec{x}_1$ to $\vec{x}_2$.

$$\|\vec{x}_1 - \vec{x}_2\|_1 = \sum_{i=1}^{3} |x_{1i} - x_{2i}| = |3 - (-2)| + |2 - 1| + |1 - 0| = 7$$

(iii) The ℓ_2-norm distance from $\vec{x}_1$ to $\vec{x}_2$.

$$\|\vec{x}_1 - \vec{x}_2\|_2 = \left(\sum_{i=1}^{3} (x_{1i} - x_{2i})^2 \right)^{1/2} = \sqrt{(3 - (-2))^2 + (2 - 1)^2 + (1 - 0)^2} = \sqrt{27}$$

(iv) The *inner* product $\vec{x}_1^T \vec{x}_2$.

$$\vec{x}_1^T \vec{x}_2 = (3)(-2) + (2)(1) + (1)(0) = -4$$

(v) The *outer* product $\vec{x}_1 \vec{x}_2^T$.

$$\vec{x}_1 \vec{x}_2^T = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} \begin{bmatrix} -2 & 1 & 0 \end{bmatrix} = \begin{bmatrix} -6 & 3 & 0 \\ -4 & 2 & 0 \\ -2 & 1 & 0 \end{bmatrix}$$

$$\mathrm{x}_1 \otimes \mathbf{x}_2 = \mathbf{x}_1 \mathbf{x}_2^T$$

$$= \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} \begin{bmatrix} -2 & 1 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 3(-2) & 3(1) & 3(0) \\ 2(-2) & 2(1) & 2(0) \\ 1(-2) & 1(1) & 1(0) \end{bmatrix}$$

$$= \begin{bmatrix} -6 & 3 & 0 \\ -4 & 2 & 0 \\ -2 & 1 & 0 \end{bmatrix}$$

Cosine of the angle between $\vec{x}_1$ and $\vec{x}_2$.

$$\begin{aligned}
\cos\theta &= \frac{\mathbf{x}_1 \cdot \mathbf{x}_2}{\|\mathbf{x}_1\| \, \|\mathbf{x}_2\|} \\
&= \frac{(3)(-2) + (2)(1) + (1)(0)}{\sqrt{3^2 + 2^2 + 1^2} \, \sqrt{(-2)^2 + 1^2 + 0^2}} \\
&= \frac{-6 + 2 + 0}{\sqrt{14} \, \sqrt{5}} \\
&= \frac{-4}{\sqrt{70}} \\
&\approx -0.478
\end{aligned}$$

The projection of $\vec{x}_1$ onto $\vec{x}_2$.

$$\frac{\mathbf{x}_1 \cdot \mathbf{x}_2}{\|\mathbf{x}_2\|^2} \, \mathbf{x}_2 = \frac{(3)(-2) + (2)(1) + (1)(0)}{(-2)^2 + 1^2 + 0^2} \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} = \frac{-6 + 2 + 0}{5} \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} = \frac{-4}{5} \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{8}{5} \\ -\frac{4}{5} \\ 0 \end{bmatrix}$$

The projection of $\vec{x}_2$ onto $\vec{x}_1$.

$$\begin{aligned}
\frac{\mathbf{x}_1 \cdot \mathbf{x}_2}{\|\mathbf{x}_1\|^2} \, \mathbf{x}_1 &= \frac{-4}{3^2 + 2^2 + 1^2} \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} \\
&= \frac{-4}{14} \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} \\
&= \begin{bmatrix} -\frac{6}{7} \\ -\frac{4}{7} \\ -\frac{2}{7} \end{bmatrix}
\end{aligned}$$

Question 2 - Multivariate data analysis

Consider the following 5 samples of some 3-dimensional data

$$\vec{x}_1 = (100, 200, 297)^T$$

$$\vec{x}_2 = (104, 202, 306)^T$$

$$\vec{x}_3 = (96, 198, 300)^T$$

$$\vec{x}_4 = (90, 195, 302)^T$$

$$\vec{x}_5 = (110, 205, 295)^T$$

For each part below, first work out the answers by hand and then check your answers using Python and numpy.

(i) The 3-D sample mean vector

$|\vec{x}|_1 = \frac{100+104+96+90+110}{5} = \frac{500}{5} = 100$

$|\vec{x}|_2 = \frac{200+202+198+195+205}{5} = \frac{1000}{5} = 200$

$|\vec{x}|_2 = \frac{297+306+300+302+295}{5} = \frac{1500}{5} = 300$

$|\vec{x}| = (100, 200, 300)^T$

(ii) The 3-D sample variance vector

$$
\begin{aligned}
s_{11} &= \frac{1}{N-1}\sum_{N}^{1}(x_{i1} - |x_1|)^2 \\
&= \frac{1}{5-1}\sum_{5}^{1}(x_{i1} - |x_1|)^2 \\
&= \frac{1}{4} \times (100-100)^2 + (104-100)^2 + (96-100)^2 + (90-100)^2 + (110-100)^2 \\
&= \frac{1}{4}(0 + 36 + 16 + 100 + 100) \\
&= \frac{1}{4}(52 + 200) \\
&= \frac{252}{4} \\
&= 63
\end{aligned}
$$

$$
\begin{aligned}
s_{22} &= \frac{1}{N-1}\sum_{N}^{1}(x_{i2} - |x_2|)^2 \\
&= \frac{1}{5-1}\sum_{5}^{1}(x_{i2} - |x_2|)^2 \\
&= \frac{1}{4} \times (200-200)^2 + (202-200)^2 + (198-200)^2 + (195-200)^2 + (205-200)^2 \\
&= \frac{1}{4}(0 + 4 + 4 + 25 + 25) \\
&= \frac{58}{4}
\end{aligned}
$$

$$
\begin{aligned}
s_{33} &= \frac{1}{N-1}\sum_{N}^{1}(x_{i3} - |x_3|)^2 \\
&= \frac{1}{5-1}\sum_{5}^{1}(x_{i3} - |x_3|)^2 \\
&= \frac{1}{4} \times (297-300)^2 + (306-300)^2 + (300-300)^2 + (302-300)^2 + (295-300)^2 \\
&= \frac{1}{4}(9 + 36 + 0 + 4 + 25) \\
&= \frac{1}{4}(74) \\
&= \frac{74}{4}
\end{aligned}
$$

This means that we have the covariance vector $x_{variance} = (\frac{252}{4}, \frac{58}{4}, \frac{74}{4})^T$

(iii) The 3-by-3 sample covariance matrix

There are another 6 elements in the s covariance matrix to calculate, however, I only need to calculate another 3 as covariance matrices are symmetrical along the diagonal.

$$\begin{aligned}
s_{12} &= \frac{1}{N-1}\sum_{N}^{1}(x_{i1} - |x_1|)(x_{i2} - |x_2|) \\
&= \frac{1}{4}\sum_{N}^{1}(x_{i1} - |x_1|)(x_{i2} - |x_2|) \\
&= \frac{1}{4} \times (0 \times 0 + 4 \times 2 + (-4) \times (-2) + (-10) \times (-5) + 10 \times 5) \\
&= \frac{1}{4} \times (8 + 8 + 50 + 50) \\
&= \frac{116}{4}
\end{aligned}$$

Now we do the same for s_{13}

$$\begin{aligned}
s_{12} &= \frac{1}{N-1}\sum_{N}^{1}(x_{i1} - |x_1|)(x_{i3} - |x_3|) \\
&= \frac{1}{4}\sum_{N}^{1}(x_{i1} - |x_1|)(x_{i3} - |x_3|) \\
&= \frac{1}{4} \times 0 \times (-3) + 4 \times 6 + (-4) \times 0 + (-10) \times 2 + 10 \times (-5) \\
&= \frac{1}{4} \times (24 - 20 - 50) \\
&= -\frac{46}{4}
\end{aligned}$$

Finally for x_{23}

$$\begin{aligned}
s_{12} &= \frac{1}{N-1}\sum_{N}^{1}(x_{i2} - |x_2|)(x_{i3} - |x_3|) \\
&= \frac{1}{4}\sum_{N}^{1}(x_{i2} - |x_2|)(x_{i3} - |x_3|) \\
&= \frac{1}{4} \times (0 \times (-3) + 2 \times 6 + (-2) \times 0 + (-5) \times 2 + 5 \times (-5)) \\
&= \frac{1}{4} \times (0 + 12 - 10 - 25) \\
&= -\frac{23}{4}
\end{aligned}$$

With all these covariances calculated, this gives us all the values we need to form the covariance matrix

$$S = \begin{bmatrix} \frac{252}{4} & \frac{116}{4} & \frac{46}{4} \\ \frac{116}{4} & \frac{58}{4} & -\frac{23}{4} \\ \frac{46}{4} & -\frac{23}{4} & \frac{74}{4} \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 252 & 116 & \frac{46}{4} \\ 116 & 58 & -\frac{23}{4} \\ 46 & 23 & \frac{74}{4} \end{bmatrix}$$

(iv) Identify the sample that has greatest ℓ_1-norm distance from the mean

We already calculated the mean vector in part (i) which is $(100, 200, 300)^T$

So now we calculate the ℓ_1-norm distance for each of the vectors

$$\begin{aligned} \|\vec{x}_1 - \bar{x}\|_1 &= |100 - 100| + |200 - 200| + |297 - 300| \\ &= 0 + 0 + 3 = 3 \end{aligned}$$

$$\begin{aligned} \|\vec{x}_2 - \bar{x}\|_1 &= |104 - 100| + |202 - 200| + |306 - 300| \\ &= 4 + 2 + 6 = 12 \end{aligned}$$

$$\begin{aligned} \|\vec{x}_3 - \bar{x}\|_1 &= |96 - 100| + |198 - 200| + |300 - 300| \\ &= 4 + 2 + 0 = 6 \end{aligned}$$

$$\begin{aligned} \|\vec{x}_4 - \bar{x}\|_1 &= |90 - 100| + |195 - 200| + |302 - 300| \\ &= 10 + 5 + 2 = 17 \end{aligned}$$

$$\begin{aligned} \|\vec{x}_5 - \bar{x}\|_1 &= |110 - 100| + |205 - 200| + |295 - 300| \\ &= 10 + 5 + 5 = 20 \end{aligned}$$

We can see that the distance is greatest to x_5

(v) Identify the sample that has greatest ℓ_2-norm distance from the mean

Now we calculate the ℓ_2 distance for each of the vectors

$$\|\vec{x}_1 - \bar{x}\|_2 = \sqrt{(100-100)^2 + (200-200)^2 + (297-300)^2} = \sqrt{0^2 + 0^2 + (-3)^2} = \sqrt{9} = 3$$

$$\|\vec{x}_2 - \bar{x}\|_2 = \sqrt{(104-100)^2 + (202-200)^2 + (306-300)^2} = \sqrt{4^2 + 2^2 + 6^2} = \sqrt{16 + 4 + 36} = \sqrt{56} \approx 7.483$$

$$\|\vec{x}_3 - \bar{x}\|_2 = \sqrt{(96-100)^2 + (198-200)^2 + (300-300)^2} = \sqrt{(-4)^2 + (-2)^2 + 0^2} = \sqrt{16 + 4 + 0} = \sqrt{20} \approx 4.472$$

$$\|\vec{x}_4 - \bar{x}\|_2 = \sqrt{(90-100)^2 + (195-200)^2 + (302-300)^2} = \sqrt{(-10)^2 + (-5)^2 + 2^2} = \sqrt{100 + 25 + 4} = \sqrt{129} \approx 11.357$$

$$\|\vec{x}_5 - \bar{x}\|_2 = \sqrt{(110-100)^2 + (205-200)^2 + (295-300)^2} = \sqrt{10^2 + 5^2 + (-5)^2} = \sqrt{100 + 25 + 25} = \sqrt{150} \approx 12.247$$

x_5 has the ℓ_2 distance to the mean.

Question 3 - Expected values

The continuous random variable X has a uniform distribution between the limits of a and b.

(i) Evaluate the expected value, $E(X)$.

(ii) Using the fact that variance is $E((X - E(X))^2)$, show that the variance of X is $\frac{1}{12}(a-b)^2$.

Question 4 - Parameter Estimation – Uniform distribution

Consider a uniform distribution between 0 and b where b is unknown, i.e. $X \sim \mathcal{U}(0, b)$.

(i) A single sample is observed with a value of 5.0. What is the maximum likelihood (ML) estimate for the parameter b?

(ii) A series of observations are observed with values $\{4.1, 4.3, 4.6, 5.0\}$. Again, what is the ML estimate for b?

(iii) After estimating b write down the expected value of X. Compare $E(X)$ with the mean value of the samples. Are they the same?

(iv) In Clockworkville buses run at precise fixed intervals 24 hours a day, i.e. a bus will arrive precisely on time every x minutes. A man, knowing this fact, but with no knowledge of the bus schedule or the interval between the buses goes to catch a bus. He has to wait exactly 5 minutes before a bus arrives. The next day a friend asks him, "How often do the buses come?" What should he reply?

2 Sheet 3

Question 1

(i) The 3-D sample mean vector