under Quiz #1.

Consider the following dummy data matrix:

4 # observation
2 # features

$$X = \begin{bmatrix} 0.5 & -1 \\ 0.5 & 1 \\ -0.5 & 1 \\ -0.5 & -1 \end{bmatrix}_{4\times 2}$$

10 I. Derive the estimated covariance matrix. Show all your work.
10 II. Find the eigen values and eigen vectors of the estimated covariance matrix. Show all your work.
5 III. Sketch the eigen vectors in the previous question on the cartesian coordinate. Show the eigen value corresponding to each eigne vector on the sketch.
5 IV. In which direction, the above data matrix has the maximum and minimum variance.
5 V. If the estimated covariance matrix multiplies the data inside the unit disk [shown below], draw [roughly] the shape of the product.

i Step 1: $\tilde{X} = X - \bar{X} = X$.

Step 2: $\hat{C} = \frac{1}{N-1} X^T . X = \frac{1}{3}\left(\underbrace{\begin{bmatrix} 0.5 & 0.5 & -0.5 & -0.5 \\ -1 & 1 & 1 & -1 \end{bmatrix}}_{2\times 4} \underbrace{\begin{bmatrix} 0.5 & -1 \\ 0.5 & 1 \\ -0.5 & 1 \\ -0.5 & -1 \end{bmatrix}}_{4\times 2} \right)$

$$= \frac{1}{3}\left(\begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}\right) = \begin{bmatrix} 1/3 & 0 \\ 0 & 4/3 \end{bmatrix} \leftarrow$$

(ii) e-value / e-vector

$$\hat{C} v = \lambda v \rightarrow |\hat{C} - \lambda \vec{I}| = 0 \qquad \vec{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\begin{vmatrix} \frac{1}{3} - \lambda & 0 \\ 0 & \frac{4}{3} - \lambda \end{vmatrix} = 0 \rightarrow (\lambda - 1/3)(\lambda - 4/3) = 0 \quad \begin{cases} \lambda_1 = 1/3 \\ \lambda_2 = 4/3 \end{cases}$$

$$v \neq 0$$

$$\hat{C} v = \lambda v \rightarrow \hat{C} v - \lambda v = 0 \rightarrow \underline{(\hat{C} - \lambda \vec{I})}\, v = 0$$

$$|\hat{C} - \lambda \vec{I}| = 0$$

$\lambda_1 = 1/3 \longrightarrow v_1 = ?$

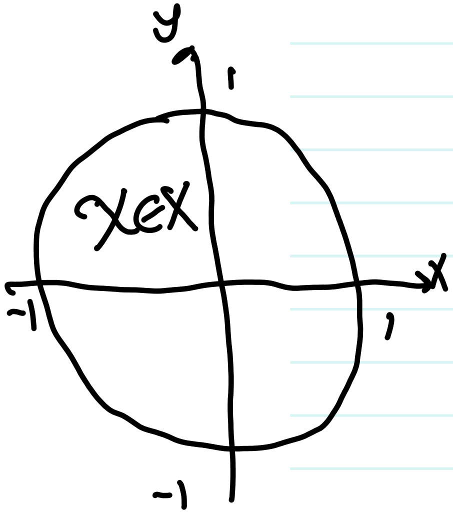
$$\hat{C} v_1 = \lambda_1 v_1 \rightarrow \begin{bmatrix} 1/3 & 0 \\ 0 & 4/3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1/3\, x \\ 1/3\, y \end{bmatrix}$$

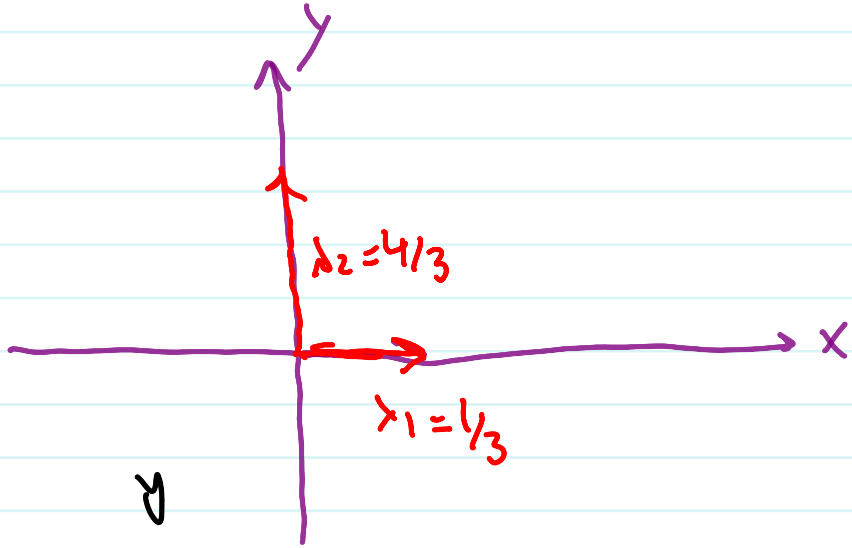
$$\begin{cases} \frac{1}{3} x = \frac{1}{3} x \\ \frac{4}{3} y = \frac{1}{3} y \end{cases} \qquad v_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$\lambda_2 = 4/3 \longrightarrow v_2 = ?$

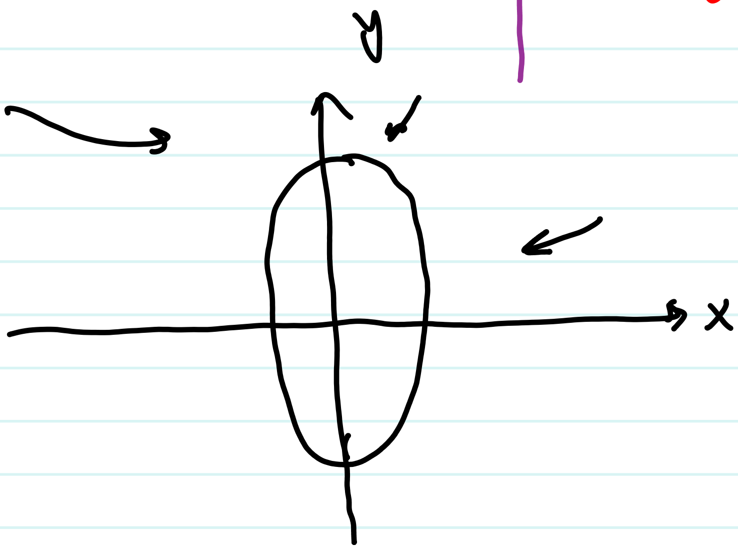
$$\hat{C} v_2 = \lambda_2 v_2 \rightarrow \begin{bmatrix} \frac{1}{3} & 0 \\ 0 & 4/3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4/3\, x \\ 4/3\, y \end{bmatrix}$$

$$\begin{cases} \frac{1}{3} x = 4/3\, x \\ 4/3\, y = 4/3\, y \end{cases} \qquad v_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

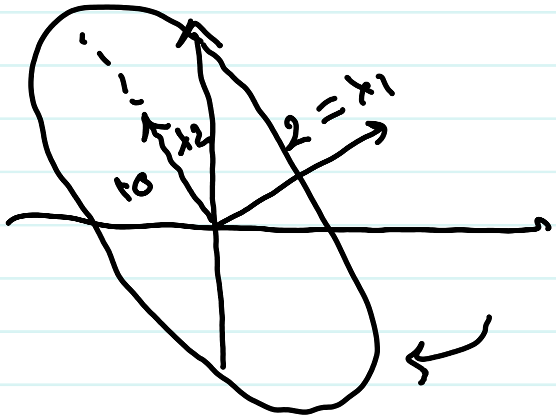




$\hat{C}\, x$



$$\hat{C}' = \begin{bmatrix} 4/3 & b \\ b & \frac{1}{3} \end{bmatrix}$$

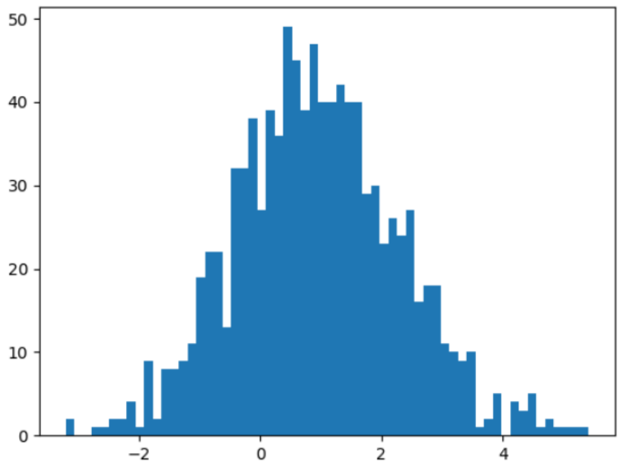


$b \neq 0 < 0$

In_Class_Sep27_pretty_table_example.py
In_Class_Sep27_VT_Data_Analytics.pdf
IQR-Standardaziation- Lecture-Boston.py
Jaccard Similarity.py
L1 and L2 Norm.pdf
LAB2.docx
LAB2.pdf
LAB_Sol.py
norm-all.py
Normalization -sklearn.py
OOP_normalization_standarziation.py
OOP_skew_call.py
Plotting exponential function.py
Rank-Example.py
relation-between-svd of X and e-values of covariance matrix.p
skewness_object01.py
skewness_object02.py
skewness_object-04.py
skewness_object_04.py

```
H =

d =
HH
for
```



Reza Jafari

nsole × Covariance-eigne-value-eigen-vector ×

```
[ 0.57663077, -0.50822058, -0.6396943 ],
[ 0.55822137, -0.32663401,  0.76269203]])
In [27]: import numpy as np
In [28]: x = np.random.normal(mean, np.sqrt(var),
In [29]: mean = 1
In [30]: var = 2
```

```
C = {ndarray: (3, 3)} [[344.74470229 257.20500807
H = {ndarray: (2, 2)} [[2.01785688 2.05222957], [2.0
HH = {DataFrame: (2, 2)} [0, 1] [0 2.017857 2.052:...
N = {int} 1000
X = {ndarray: (1000, 3)} [[8.15389049e-01 2.255758
X_tid = {DataFrame: (10000, 2)} [0, 1] [0 1.15059 ...
a = {ndarray: (3,)} [825.41207595 79.0593311 86.6
b = {ndarray: (3, 3)} [[ 0.5965617 0.79688272 -0.09
d = {PrettyTable} +--------------------+-----------
i = {int} 1
mean = {int} 1
s = {ndarray: (3,)} [28.72998566 9.30839748 8.891
u = {ndarray: (1000, 1000)} [[-2.76158847e-02 -2.11
v = {ndarray: (3, 3)} [[-0.5965617 -0.57663077 -0.5
var = {int} 2
vh = {ndarray: (2, 2)} [[ 0.44395552 0.89604882], [
x = {ndarray: (1000,)} [-7.20016292e-01 -4.1920559
```

likihood function prior

$$\underbrace{P(Y \mid x_1, x_2, \cdots x_d)}_{\text{posterior}} = \frac{\overbrace{P(x_1, x_2, \cdots, x_d \mid Y)}\,\overbrace{P(Y)}}{P(x_1, \ldots x_d)}$$

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{\overbrace{P(A \mid B)} \cdot P(B)}{P(A)}$$

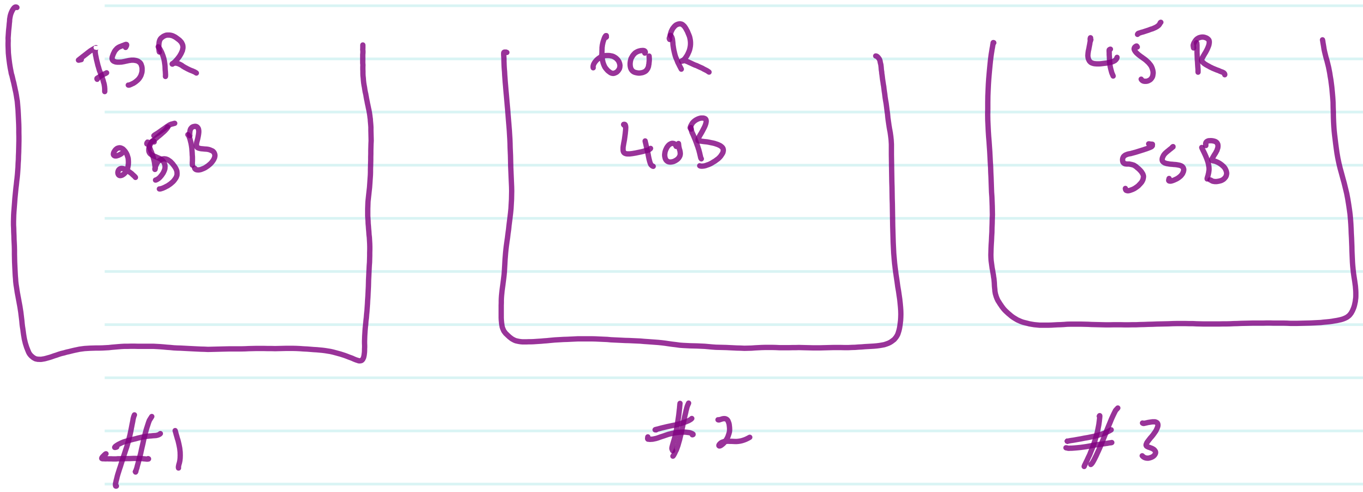
$$P(\text{cancer} \mid x_1, \ldots x_d) < P(\overline{\text{cancer}}) \mid x_1, \ldots x_d)$$

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} \underset{A, B}{\overset{\text{indepdent}}{=}} \frac{\cancel{P(A)} \cdot P(B)}{P(\cancel{A})} = P(B)$$

$$\boxed{f_{X \mid Y}(\lambda_1 \mid \lambda_2) = f_X(\lambda_1)}$$

- **Example**: There are three bags that each contain 100 marbles:
  - Bag 1 has 75 red and 25 blue marbles;
  - Bag 2 has 60 red and 40 blue marbles;
  - Bag 3 has 45 red and 55 blue marbles;

  I choose one of the bags at random and then pick a marble from the chosen bag, also at random. What is the probability that the chosen marble is red?

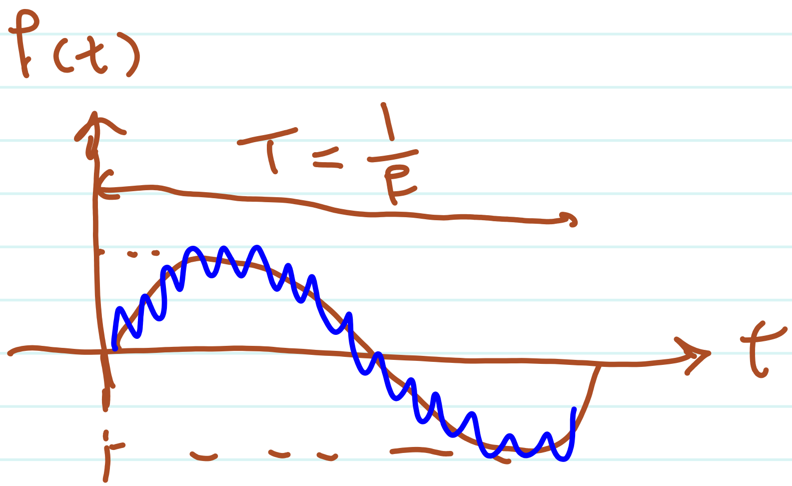


$$P(\text{red}) = \frac{1}{3}\left(\frac{75}{100} + \frac{60}{100} + \frac{45}{100}\right) = \frac{\overset{6}{\cancel{18}}0}{100}\cdot\frac{1}{\cancel{3}} = 60\%$$

$$P(\text{Stock}\uparrow \mid \text{Lanched}) > P(\text{Stock}\uparrow \mid \overline{\text{Launch}})$$

$$f(t) = A\sin(\omega t + \theta)$$

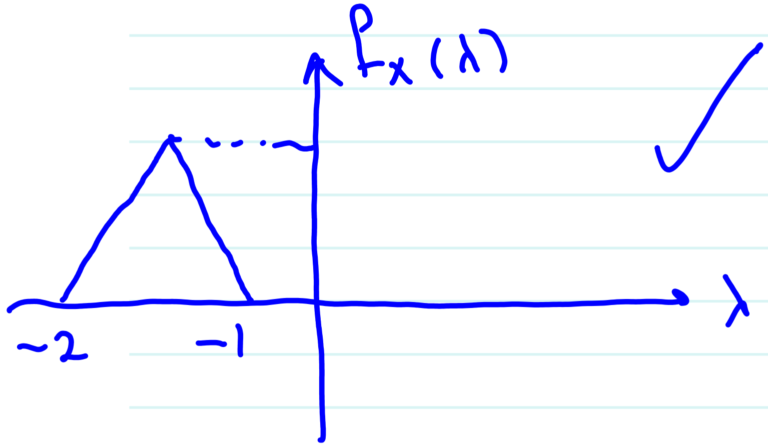
$$\left.\begin{array}{l} A \\ \omega \\ \theta = 0 \end{array}\right\} \text{fixed}$$

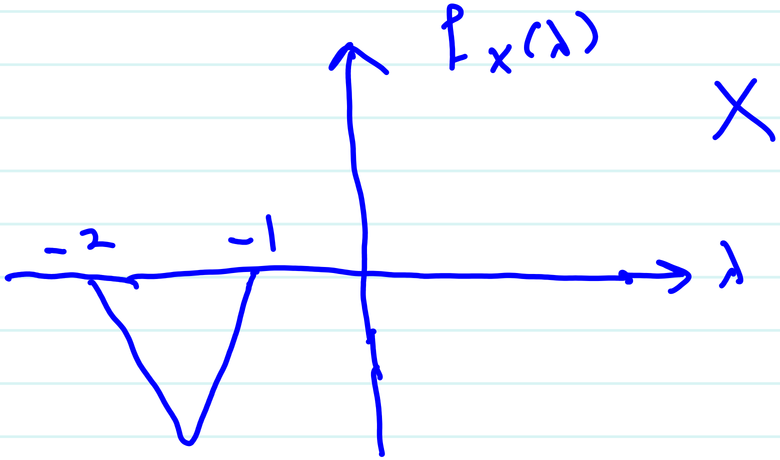


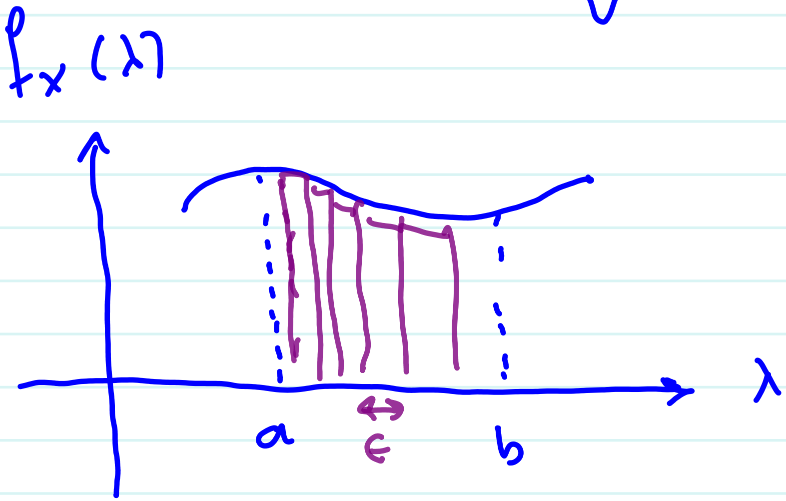
deterministic process

$$f'(t) = f(t) + \epsilon$$

$$\epsilon \sim \mathcal{N}(0, \sigma_e^2)$$





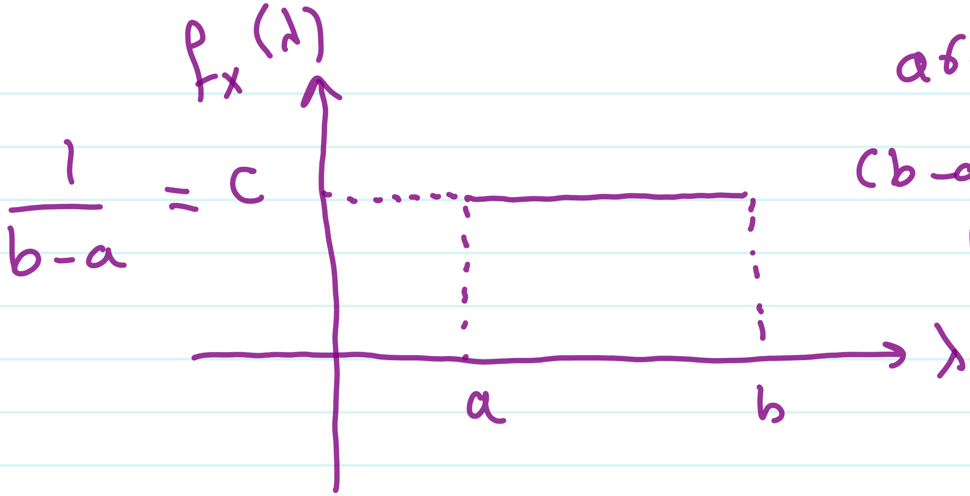


$$\Pr(a < X < b) = \int_a^b f_X(\lambda)\, d\lambda$$

Let consider a uniform density function for a random variable X. This random variable is said to have *Uniform(a,b)* distribution. Or it can mathematically to be written as:

$$f_X(\lambda) = \begin{cases} c, & 0 < a < \lambda < b \\ 0, & \text{Else} \end{cases}$$

Graph the above density function and find c in terms if a and b.



area = 1

$(b-a) \cdot c = 1$

$\boxed{c = \frac{1}{b-a}}$

$$\int_{-\infty}^{\infty} f_X(\lambda)\,d\lambda = 1 \rightarrow \int_a^b c\,d\lambda = 1 \rightarrow c\lambda\Big|_a^b = 1$$

$c(b-a)=1 \rightarrow \boxed{c = \frac{1}{b-a}}$

Let X be a continuous random variable with the following pdf:

$e^{-\infty} = \frac{1}{e^{\infty}} = 0$

$$f_X(\lambda) = \begin{cases} ce^{-\lambda}, & \lambda \geq 0 \\ 0, & \text{Else} \end{cases}$$

$\int ae^{-cx}\,dx = \frac{ae^{-cx}}{-c}$

$\int ax^n\,dx = \frac{ax^{n+1}}{n+1}$

$\frac{d(ax^n)}{dx} = anx^{n-1}$

1. Find c.
2. Find the probability distribution function $F_X(\lambda)$
3. Find $P(1 < X \leq 3)$
4. Find $P(X = 2)$
5. Find $P(X \in [0, 1] \cup [3, 4])$

$\int_{-\infty}^{\infty} f_X(\lambda)\,d\lambda = 1 \rightarrow \int_0^{\infty} ce^{-\lambda}\,d\lambda = 1$

$\frac{ce^{-\lambda}}{-1}\Big|_0^{\infty} = 1 \rightarrow -c(e^{-\infty} - e^{0}) = 1 \rightarrow -c(-1) = 1$

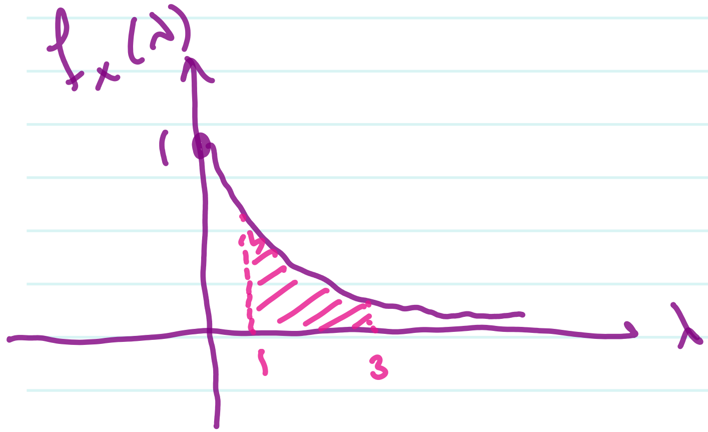
$\boxed{c = 1}$

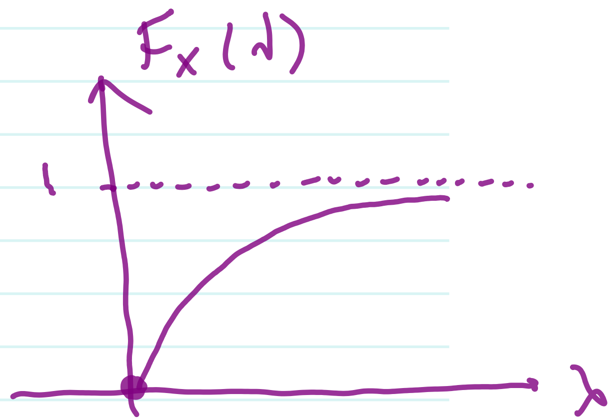
(ii) $F_X(\lambda) = \int_{-\infty}^{\lambda} f_X(\mu)\, d\mu$

$$= \int_0^{\lambda} e^{-\mu}\, d\mu = \left. \frac{e^{-\mu}}{-1} \right|_0^{\lambda} = -(e^{-\lambda} - e^{0}) = 1 - e^{-\lambda}$$

$f_X(\lambda) = e^{-\lambda} \quad : \quad \lambda \geq 0$

$F_X(\lambda) = 1 - e^{-\lambda} \quad : \quad \lambda \geq 0$





$Pr(1 < X < 3) = \int_1^3 e^{-\lambda}\, d\lambda$

$$= \frac{e^{-\lambda}}{-1}\Big|_1^3 = -(e^{-3} - e^{-1}) = \boxed{e^{-1} - e^{-3}}$$

$$Pr(X=2) = \lim_{\epsilon \to 0} Pr(2-\epsilon < X < 2+\epsilon)$$

$$= \lim_{\epsilon \to 0} F_X(2+\epsilon) - \lim_{\epsilon \to 0} F_X(2-\epsilon) = 0$$

$$Pr(X \in [0,1] \cup [3,4]) = \int_0^1 e^{-\lambda}\, d\lambda + \int_3^4 e^{-\lambda}\, d\lambda$$

$$= \frac{e^{-\lambda}}{-1}\Big|_0^1 + \frac{e^{-\lambda}}{-1}\Big|_3^4 =$$

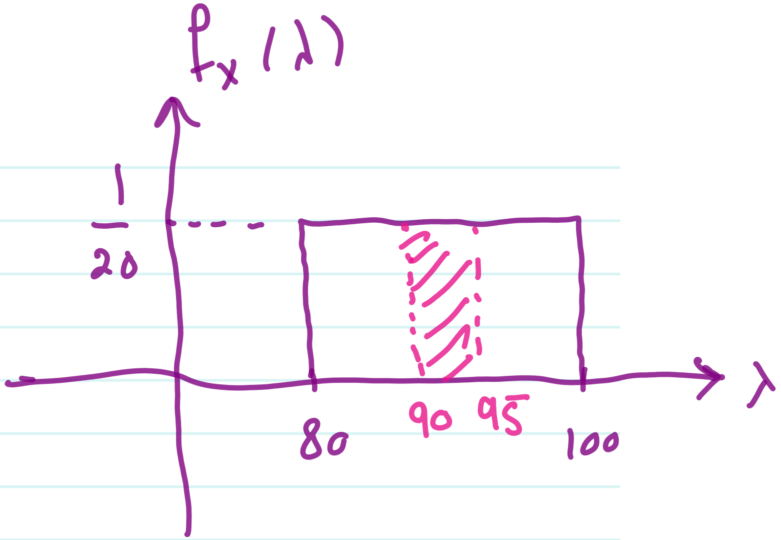
$$= -(e^{-1} - 1) - (e^{-4} - e^{-3}) =$$

$$= \boxed{1 - e^{-1} - e^{-4} + e^{-3}}$$

Let X be a continuous random variable that is equally likely to be any valye between 80 and 100.

1. Graph the corresponding probability density function.
2. Find $P(90 < X \leq 95)$

$$f_X(x) = \begin{cases} \frac{1}{20} & : 80 < x < 100 \\ 0 & : \text{Else} \end{cases}$$



$$Pr(90 < X < 95) = 5 \cdot \frac{1}{20} = 25\%$$

Let X and Y be two jointly continuous random variables (uniformly distributed) with joint probability density function:

$$f_{XY}(\lambda_1, \lambda_2) = \begin{cases} \lambda_1 + \overset{3/2}{c}\lambda_2^2, & 0 \le \lambda_1 \le 1, 0 \le \lambda_2 \le 1 \\ 0, & \text{Else} \end{cases}$$

1. Find c.
2. Find $P(0 < X \le \frac{1}{2}, 0 < Y \le \frac{1}{2})$

λ_2 is changing

$$\overbrace{\int_{-\infty}^{\infty} \underbrace{\int_{-\infty}^{\infty} f_{xy}(\lambda_1, \lambda_2)\, d\lambda_1}_{\substack{\lambda_1 \text{ is changing} \\ \lambda_2 \text{ is fixed}}} d\lambda_2}^{} = 1 \rightarrow \int_0^1 \underbrace{\int_0^1 (\lambda_1 + c\lambda_2^2)\, d\lambda_1}_{} d\lambda_2 = 1$$

$$\rightarrow \int_0^1 \left(\frac{\lambda_1^2}{2} + c\lambda_1\lambda_2^2\right)\Big|_0^1 d\lambda_2 = 1$$

$$\int_0^1 \left(\frac{1}{2} + c\lambda_2^2\right) d\lambda_2 = 1$$

$$\left(\frac{\lambda_2}{2} + \frac{c\,\lambda_2^3}{3}\right)\Bigg|_0^1 = 1 \rightarrow \frac{1}{2} + \frac{c}{3} = 1$$

$$\frac{c}{3} = \frac{1}{2} \rightarrow \boxed{c = \frac{3}{2}}$$

$$P_r\left(0 < X \le \tfrac{1}{2},\ 0 < Y \le \tfrac{1}{2}\right) = \int_0^{\frac{1}{2}}\int_0^{\frac{1}{2}} \left(\lambda_1 + \frac{3}{2}\lambda_2^2\right) d\lambda_1\, d\lambda_2$$

$$= \int_0^{\frac{1}{2}} \left(\frac{\lambda_1^2}{2} + \frac{3}{2}\lambda_1\,\lambda_2^2\right)\Bigg|_0^{\frac{1}{2}} d\lambda_2 = \int_0^{\frac{1}{2}} \left(\frac{1}{8} + \frac{3}{2}\,\frac{1}{2}\cdot\lambda_2^2\right) d\lambda_2$$

$$= \left(\frac{\lambda_2}{8} + \frac{\cancel{3}}{4}\cdot\frac{\lambda_2^3}{\cancel{3}}\right)\Bigg|_0^{\frac{1}{2}} = \frac{1}{16} + \frac{1}{32} = \boxed{\frac{3}{32}}$$

Let X and Y be two jointly continuous random variables with the joint probability density function:

$$f_{XY}(\lambda_1,\lambda_2)=\begin{cases} c\lambda_1^2+\frac{\lambda_1\lambda_2}{3}, & 0\le\lambda_1\le 1, 0\le\lambda_2\le 2\\ 0, & \text{Else}\end{cases}$$

1. Find c.
2. Find $P(X+Y\ge 1)$

1

ϕ

$\int\int$

ϕ

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(c\lambda_1^2+\frac{\lambda_1\lambda_2}{3}\right)d\lambda_1\,d\lambda_2=1$$

$$\int_0^1\int_0^2\left(c\lambda_1^2+\frac{\lambda_1\lambda_2}{3}\right)d\lambda_2\,d\lambda_1=1$$

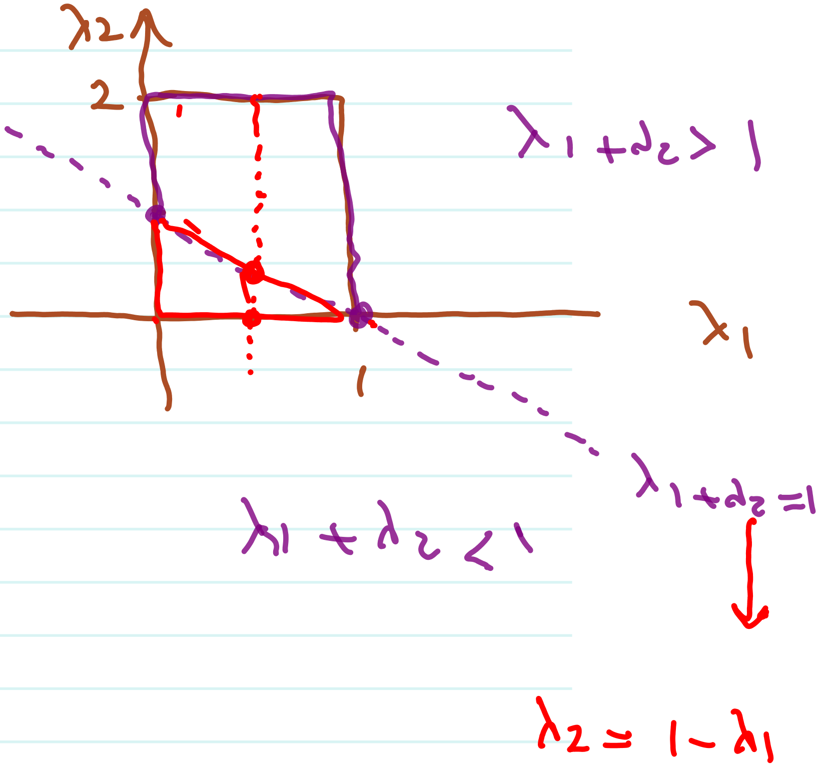
$$\int_0^1\left(c\lambda_2\lambda_1^2+\frac{\lambda_2^2\lambda_1}{6}\right)\Big|_0^2=1$$

$$\int_0^1\left(2c\lambda_1^2+\frac{2}{3}\lambda_1\right)d\lambda_1=1$$

$$\left(\frac{2c\lambda_1^3}{3} + \frac{2}{3}\cdot\frac{\lambda_1^2}{2}\right)\Big|_0^1 = 1$$

$$\frac{2c}{3} + \frac{1}{3} = 1 \rightarrow \frac{2c}{3} = \frac{2}{3} \rightarrow \boxed{c = 1}$$

$$\underline{\Pr(\lambda_1 + \lambda_2 \geq 1)}$$



$$= 1 - \Pr(\lambda_1 + \lambda_2 < 1)$$

$$= 1 - \int_0^1 \int_0^{1-\lambda_1} \left(\lambda_1^2 + \frac{\lambda_1\lambda_2}{3}\right) d\lambda_2 \; d\lambda_1$$

$$= 1 - \int_0^1 \left(\lambda_2\lambda_1^2 + \frac{\lambda_1\lambda_2^2}{6}\right)\Big|_0^{1-\lambda_1} d\lambda_1$$

$$1-2\lambda_1+\lambda_1^2$$

$$= 1 - \int_0^1 \left(\lambda_1^2 (1-\lambda_1) + \frac{\lambda_1 (1-\lambda_1)^2}{6} \right) d\lambda_1$$

$$= 1 - \int_0^1 \left(\lambda_1^2 - \lambda_1^3 + \frac{\lambda_1}{6} - \frac{1}{3}\lambda_1^2 + \frac{\lambda_1^3}{6} \right) d\lambda_1$$

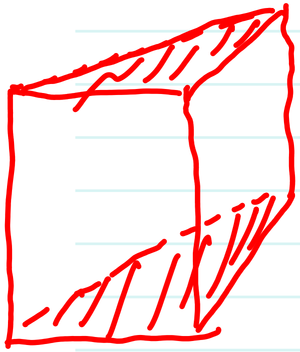
$$= 1 - \left(\frac{\lambda_1^3}{3} - \frac{\lambda_1^4}{4} + \frac{\lambda_1^2}{12} - \frac{\lambda_1^3}{9} + \frac{\lambda_1^4}{24} \right) \Big|_0^1$$

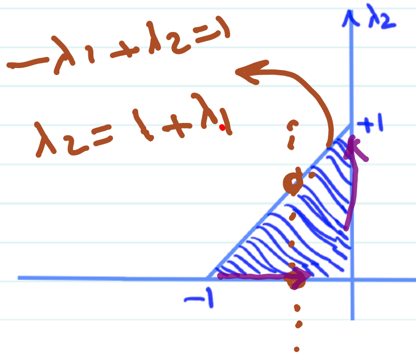
$$= 1 - \left(\underbrace{\frac{1}{3} - \frac{1}{4}}_{\frac{1}{12}} + \frac{1}{12} - \frac{1}{9} + \frac{1}{24} \right) = 1 - \left(\frac{1}{6} - \frac{1}{9} + \frac{1}{24} \right)$$

$$= \boxed{1 - \left(\frac{5}{24} - \frac{1}{9} \right) =}$$

- Let consider two random variables X and Y with the following joint probability density function:

$$f_{XY}(\lambda_1, \lambda_2) = \begin{cases} \overset{2}{c}, & -1 \le \lambda_1 \le 0, 0 \le \lambda_2 \le 1 \\ 0, & \text{Else} \end{cases}$$





Volum = 1

area * height = 1

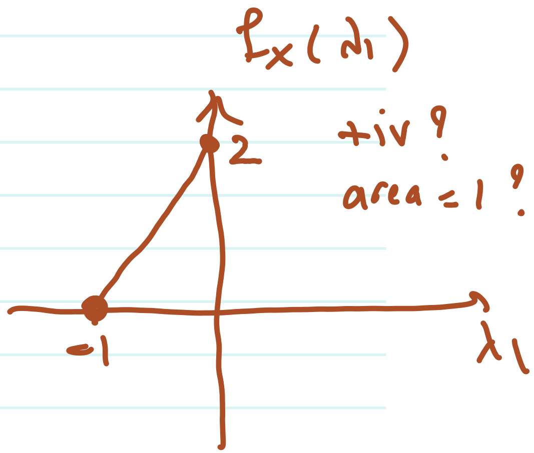
$\frac{1}{2} \cdot c = 1$

$\boxed{c = 2}$

$$f_X(\lambda_1) = \int_{-\infty}^{\infty} f_{x,y}(\lambda_1, \lambda_2)\, d\lambda_2$$

$$= \int_0^{1+\lambda_1} 2\, d\lambda_2 = 2(1+\lambda_1)$$

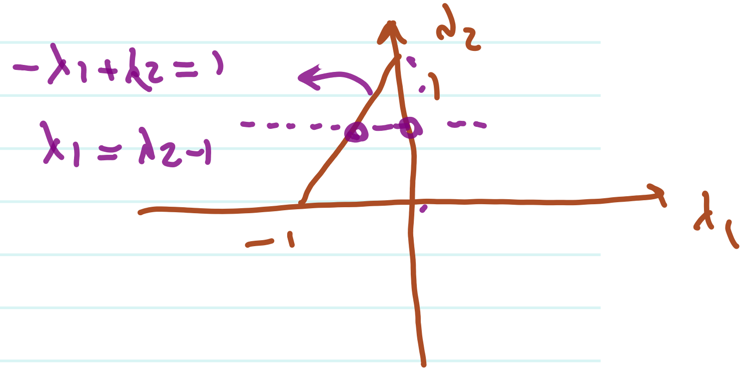
$$f_X(\lambda_1) = \begin{cases} 2(1+\lambda_1) & : \quad -1 \le \lambda_1 < 0 \\ 0 & : \quad \text{Else} \end{cases}$$



+iv ?

area = 1 ?

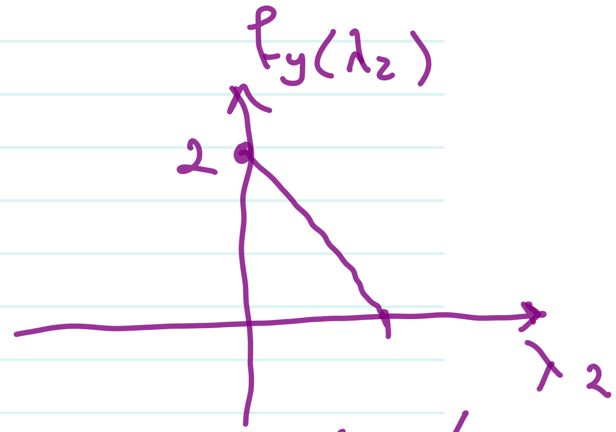
$$f_y(\lambda_2) = \int_{-\infty}^{\infty} f_{xy}(\lambda_1, \lambda_2)\, d\lambda_1$$



$$= \int_{\lambda_2 - 1}^{0} 2 \quad d\lambda_1$$

$$= 2(1 - \lambda_2)$$

$$f_y(\lambda_2) = \begin{cases} 2(1-\lambda_2) & : \quad 0 < \lambda_2 < 1 \\ 0 & : \quad \text{Else} \end{cases}$$



Valid: +ive ✓

area = 1

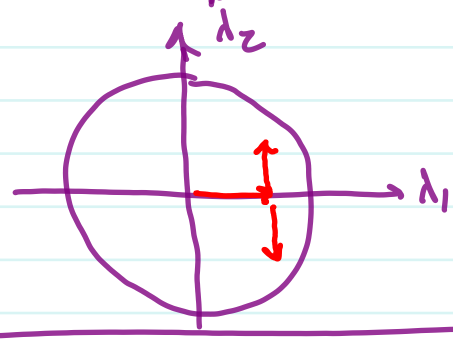
---

independancy :

$$f_{xy}(\lambda_1, \lambda_2) \stackrel{!}{=} f_x(\lambda_1) \cdot f_y(\lambda_2) \quad \forall \lambda_1, \lambda_2$$

$$2 \neq 2(1+\lambda_1)\, 2(1-\lambda_2)$$

X, y are not independent = dependent.



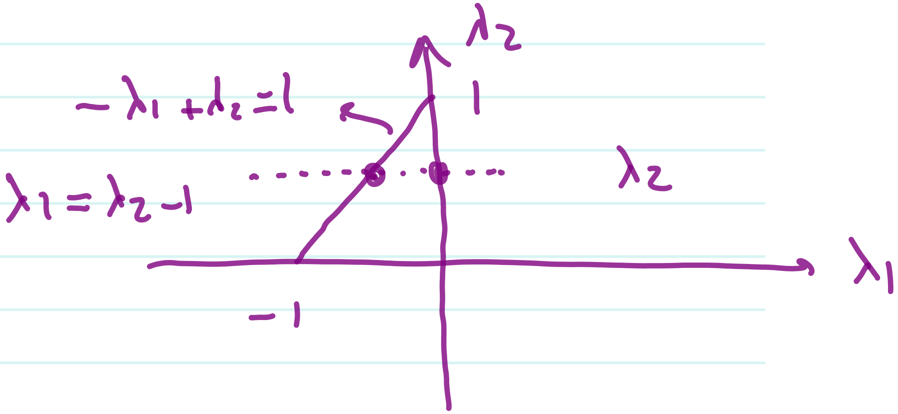
$$f_X(\lambda_1) = \begin{cases} 2(1+\lambda_1) : & -1 < \lambda_1 < 0 \\ 0 & : \text{else} \end{cases}$$

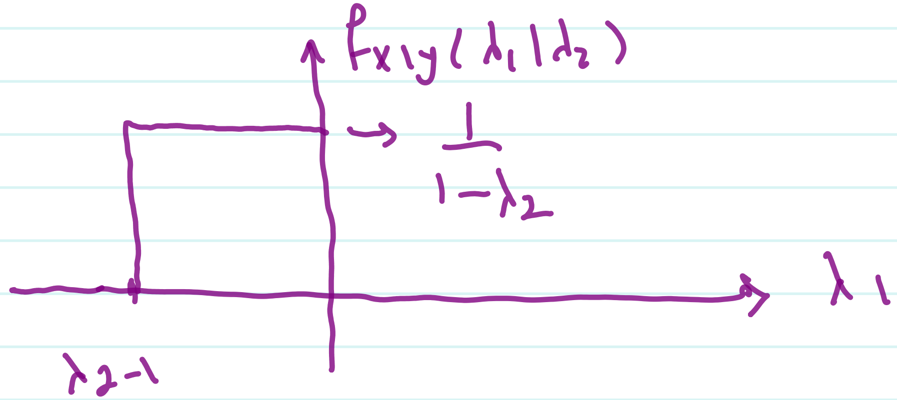
$$f_Y(\lambda_2) = \begin{cases} 2(1-\lambda_2) : & 0 < \lambda_2 < 1 \\ 0 & : \text{Else} \end{cases}$$

$$f_{X|Y}(\lambda_1 \mid \lambda_2) = \frac{f_{X,Y}(\lambda_1, \lambda_2)}{f_Y(\lambda_2)} = \frac{2}{2(1-\lambda_2)} = \frac{1}{1-\lambda_2}$$

$$f_{X|Y}(\lambda_1 \mid \lambda_2 = \tfrac{1}{4}) = \frac{1}{1-\frac{1}{4}} = \frac{4}{3}$$

$$f_{X|Y}(\lambda_1|\lambda_2) = \frac{1}{1-\lambda_2}$$

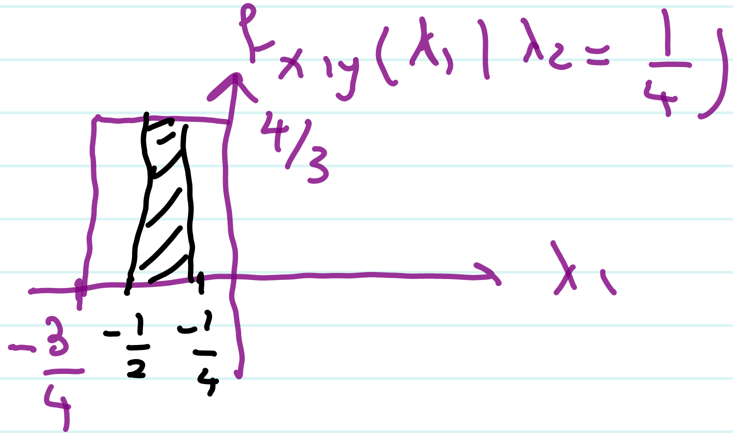




valid = +ive

area = 1

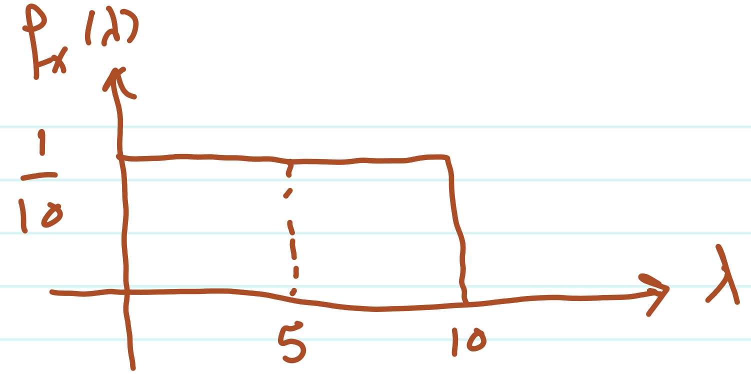
$$f_{X|Y}(x|\lambda_2=\tfrac{1}{4}) = \frac{4}{3}$$



$$Pr\left(-\frac{1}{2} < \lambda_1 < -\frac{1}{4} \,\middle|\, \lambda_2 = \frac{1}{4}\right) = \frac{4}{3} * \frac{1}{4} = \frac{1}{3}$$

$$\sigma_{xy} = E[(x-\mu_x)(y-\mu_y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (\lambda_1-\mu_x)(\lambda_2-\mu_y)\, f_{xy}(\lambda_1,\lambda_2)\, d\lambda_1\, d\lambda_2$$

- Let consider a continuous random variable X to be defined uniformly between 0 and 10. What is the mean of X and what is the variance?
- **Answer** : $\mu_x = 5$ and $\sigma_x^2 = 8.3$



$$f_x(\lambda) = \begin{cases} \frac{1}{10} : & 0 < \lambda < 10 \\ 0 : & \text{Else} \end{cases}$$

$$E[x] = \int_{-\infty}^{\infty} \lambda f_x(\lambda)\, d\lambda = \int_0^{10} \frac{\lambda}{10}\, d\lambda = \left.\frac{\lambda^2}{20}\right|_0^{10}$$

$$= \frac{100}{20} = 5$$

$$\sigma_x^2 = E[(x-5)^2] = \int_0^{10} \frac{(\lambda-5)^2}{10}\, d\lambda$$

$$= \frac{1}{10}\int_0^{10} (\lambda^2 - 10\lambda + 25)\, d\lambda = \frac{1}{10}\left(\frac{\lambda^3}{3} - 5\lambda^2 + 25\lambda\right)\Big|_0^{8}$$

$$= \frac{1}{1\cancel{0}}\left(\frac{100\cancel{0}}{3} - \overbrace{500+250}^{-25\cancel{0}}\right) = \frac{100}{3} - 25 = \frac{25}{3} \simeq \boxed{8.3}$$